

OM of: US-09-236-995D-2 to: GenEmbl.* out_format : pfs
Date: Mar 9, 2002 7:40 PM
About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
-MODEL=framet_p2n.model -DEV=xlp
-Q/cgnt2_1/USPTO.spool/US09236995/runat_08032002_113149_25104/app_query.fasta_1.1273
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPCL=0.100 -LOOPEXT=0.000 -QGAPOP=4.500
-QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-DELEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-DELEXT=7.000 -START=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0
-MAXLEN=200000000 -USER=US09236995 -CGN1_1.8116 -NCP0=6
-ICPU=3 -LONGLOG -NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-09-236-995D-2
Query length: 982
Database: GenEmbl.*
Database sequences: 1472140
Database length: 341344837
Search time (sec): 3984.260000

score_list:	Strd Orig	ZScore	EScore	Len	Documentation
gb_pl:AF093627	+ 4742.00	6447.72	0.0	3285	! AF093627 Zea mays poly(ADP-ribose)
gb_pl:ZMPARP2	+ 4666.50	6344.95	0.0	3211	! AJ222589 Zea mays mRNA for poly(ADP-ribose)
gb_pl:ATH131705	+ 2774.50	3764.21	2.2e-201	3187	! AJ131705 Arabidopsis thaliana
gb_pl:AC006593	- 1973.50	2640.95	7.9e-139	79663	! AC006593 Arabidopsis thaliana
gb_ro:AF168781	+ 1602.50	2165.95	2.3e-112	3047	! AF168781 Arabidopsis thaliana
gb_pat:AS2134	+ 1599.50	2159.77	5.0e-112	3792	! AS2134 Sequence 1 from Patent
gb_ro:HUMPOLP	+ 1597.00	2156.75	7.4e-112	3640	! M1812 Human poly(ADP-ribose)
gb_ov:GGADPRP	+ 1594.50	2155.07	9.2e-112	3036	! X52670 Chicken mRNA for poly(ADP-ribose)
gb_ro:HUMRISDAD	+ 1593.00	2150.90	1.6e-111	3795	! J03473 Human poly(ADP-ribose)
gb_ro:BC012041	+ 1593.00	2150.78	1.6e-111	3845	! BC012041 Mus musculus, Simlani
gb_ro:RN094340	+ 1592.00	2151.38	1.5e-111	3128	! U94340 Rattus norvegicus poly(ADP-ribose)
gb_pat:IL14359	+ 1587.00	2142.84	4.4e-111	3747	! I14359 Sequence 5 from patent
gb_ro:WMADPRP	+ 1584.50	2141.01	5.6e-111	3172	! I14206 Mouse mRNA for poly(ADP-ribose)
gb_pat:AX058340	+ 1584.00	2140.72	5.8e-111	3045	! AX058340 Sequence 24 from patent
gb_pat:AX062277	+ 1584.00	2140.72	5.8e-111	3045	! AX062277 Sequence 136 from patent
gb_ro:HUMPPOL	+ 1584.00	2138.97	7.3e-111	3660	! M32721 Human poly(ADP-ribose)
gb_ov:KELFADPRP	+ 1576.00	2128.24	2.9e-110	3589	! D14567 Frog mRNA for poly(ADP-ribose)
gb_ov:BOVPAS	+ 1548.00	2089.98	3.9e-108	3617	! D12139 X. laevis PARP gene encoding poly(ADP-ribose)
gb_ov:XLPARG	+ 1518.50	2050.90	5.8e-106	3200	! AX058362 Sequence 46 from patent
gb_pat:AX058362	+ 1381.00	1862.59	1.8e-95	3463	! D1482 Sarcophaga peregrina mRNA for poly(ADP-ribose)
gb_in:SPEDPRP	+ 1285.50	1696.13	3.4e-86	155026	! AL442128 Human DNA sequence
gb_ro:AL442128	+ 1269.50	1711.61	4.6e-87	3080	! D13806 Fruit fly mRNA for poly(ADP-ribose)
gb_in:DROADPRP	+ 1243.50	1670.14	9.4e-85	5787	! L24094 Gorilla gorilla ADP-ribose
gb_ro:CORADPRP	+ 1237.00	1662.03	2.7e-84	5345	! I14360 Sequence 7 from patent
gb_pat:IL14360	+ 1237.00	1656.81	5.2e-84	9254	! L14752 Homo sapiens poly(ADP-ribose)
gb_ro:HUMADPRP2	+ 1002.00	1351.99	5.0e-67	1771	! M17081 Human placental poly(ADP-ribose)
gb_ro:AF126717	+ 993.50	1340.61	2.1e-66	1732	! AF126717 Mus musculus mPARP mRNA
gb_ro:RNPARP2	+ 990.50	1335.44	4.1e-66	1939	! X65497 R. norvegicus mRNA for poly(ADP-ribose)
gb_pr:S54638	+ 979.50	1317.35	4.2e-65	2682	! S54638 [pseudogene] (A)-poly(ADP-ribose)
gb_pat:IL14357	+ 948.50	1275.06	9.6e-63	2682	! I14357 Sequence 3 from patent
gb_pat:AX062218	+ 898.00	1204.18	8.5e-59	3308	! AX062218 Sequence 177 from patent
gb_pr:HS236912	+ 851.50	1146.79	1.3e-55	1754	! AJ236912 Homo sapiens mRNA for poly(ADP-ribose)
gb_ro:HS236876	+ 851.50	1146.56	1.4e-55	1797	! AJ236876 Homo sapiens mRNA for poly(ADP-ribose)
gb_pat:AX058317	+ 851.50	1146.47	1.4e-55	1814	! AX058317 Sequence 1 from patent
gb_pat:AX008892	+ 851.50	1146.32	1.4e-55	1843	! AX008892 Sequence 1 from patent
gb_pat:AX058374	+ 851.50	1146.16	1.4e-55	1874	! AX058374 Sequence 58 from patent
gb_pat:AX058374	+ 851.50	1144.87	1.7e-55	2147	! Z48243 A. thaliana PARP mRNA for poly(ADP-ribose)
gb_pl:ATPARP	+ 842.50	1133.71	7.2e-55	1910	! AF085734 Homo sapiens NAD+ ADP-ribose
gb_ro:AF085734	+ 842.50	1133.71	7.2e-55	1910	! AF085734 Homo sapiens NAD+ ADP-ribose
gb_pat:AX058338	+ 840.00	1131.36	9.7e-55	1707	! AX058338 Sequence 22 from patent
gb_ro:MMU7780	+ 840.00	1131.36	9.7e-55	1707	! AJ007780 Mus musculus mRNA for poly(ADP-ribose)

gb_ro:AF072521 + 840.00 1131.04 1.0e-54 1766 ! AF072521 Mus musculus poly(ADP-ribose)
gb_pl:ZMPARP1 + 838.50 1126.50 1.8e-54 2295 ! AJ222588 Zea mays mRNA for poly(ADP-ribose)
gb_pr:AK001980 + 838.00 1127.22 1.6e-54 1980 ! AK001980 Homo sapiens cDNA
gb_in:DMPARP5 + 830.50 1117.79 5.5e-54 1822 ! AF051548 Drosophila melanogaster

seq_name: gb_pl:AF093627

seq_documentation_block:
LOCUS AF093627 3285 bp mRNA PLN 29-NOV-1998
DEFINITION Zea mays poly(ADP-ribose) polymerase (PARP1) mRNA, complete cds.

ACCESSION AF093627 GI:3928870

VERSION AF093627.1

KEYWORDS Zea mays.

SOURCE Zea mays.

ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC

REFERENCE 1 (bases 1 to 3285)

AUTHORS Mahajan,P.B. and Zuo,Z.

TITLE Purification and cDNA cloning of maize Poly(ADP-ribose) polymerase

JOURNAL Plant Physiol. 118 (3), 895-905 (1998)

MEDLINE 99026291

REFERENCE 2 (bases 1 to 3285)

AUTHORS Mahajan,P.B. and Zuo,Z.

TITLE Direct Submission

JOURNAL Submitted (22-SEP-1998) Crop Protection, Pioneer HiBred

International, Inc., 7300 NW 62nd Avenue, Johnston, IA 50131-1004,

USA

FEATURES

source

1..3285 Location/Qualifiers

/organism="Zea mays"

/db_xref="taxon:4577"

1..3285

/gene="PARP1"

100..3042

/gene="PARP1"

/EC_number="2.4.2.30"

/function="add ADP-ribose"

/codon_start=1

/product="poly(ADP-ribose) polymerase"

/protein_id="AAC79704.1"

/db_xref="GI:3928871"

/translation="MAAPPAKAAEAKASRASCSCRSPIAKDOLRLGKMWAOASOF
GFPMNHAHCIFSKNQIKSVDDDEGIDALRWDDQEKIRNYVGSASAGTSSFAAPPE
KCTIETASRTSCRCSEKIKGVSRLSAKLESGKGGKIPWTHANCFEVSFAATPE
KFGWDTLSDDEKRTLDLVKQVQNEQKSKKSKSENDSIDSKSARLDESGT
VRNKGQVLDPRGNTSADQLKQKSDTLKMLKDLKTHVSAELRLMLEANGDT
SGPERHLIDRCADGMFLGALGPCVCGANGMYVNGOYCGNVSSEKCTYSATEPVR
VKKWQIPHTKNDYLMKWKSKQKVKPKPVLPPMSPEKSGSKATORTSILSSKGLDK
LRFVSQGSKEAANEITEKULKAGAFYARVKDIDCLACGLDNEAENAEVRARLK
IPIVREGYIGECVKKNNKMPFDLYKLENALSSKSTVTIVKVGSAVHSSGLQDTA
HILEDKSIYNATLNNSDALGVNSYVLIQIIEQDQSECVYKRWKGVSEKIGSQK
LEMSKTEALKEKRLFLKKTGNSWEAKETNFKPKPGFYPLDVGKVPKPKRDE
ISEMKSALPQLLEKMLFNVTYRAAMFEETINMSPEKLSKENIKGPEALTE
IONLLKDTAQALAVRESLVAASNFEETLIPSHIPIINDEDDLMKAKMLBALQDI
EIAKIVGSDSDSELDLKKMLKHCIDITPLANDSEYKLIIEYLNTHTAPTHKMSL
ELEVEFSLDGLSKSYKNNLHNKMLHNSRLTNFVGLISQSLRGLTAPPEAPVTG
YMFKGILYFADLVSKAQYCYVDNRNPNVGLMLLSEVALGDMYELKATSMKPPKGRK
STKGLGKTVPLESEFVKWRDDVVVPCGKPPSPSSIRSELMIYNEYIVNTSQVKMQFL
KVRPHKR"

alignment_scores:

Quality: 4742.00 Length: 985

Ratio: 5.002 Caps: 5

Percent Similarity: 96.244 Percent Identity: 95.635

alignment_block:

US-09-236-995D-2 x AF093627

Align seg 1/1 to: AF093627 from: 1 to: 3285

1 MetAlaAlaProProLysAlaTrpLysAlaGluTyrAlaLysSerGlyAr 17
100 ATGGCGCGCGCCCAAGAGCGGTGAAGCGGGAGTATGCCAAGTCTGGGCG 149
17 gAlaSerCysLysSerCysArgSerProIleAlaLysAspGlnLeuArgL 34
150 GGCTCTGTCAGTCAATGCGCGGTCCCTATCGCCAGGACCAGCTCCGTC 199
34 euGlyLysMetValGlnAlaSerGlnPheAspGlyPheMetProMetTrp 50
200 TTGGCAAGATGTTACAGCGCTCACAGTTCGACGGCTTCATGCCGATGG 249
51 AsnHisAlaArgCysIlePheSerLysLysAsnGlnIleLysSerValAs 67
250 AACCATGCCAGGTGCATCTTCAGCAGAGAACACGATAAATCCGTGA 299
67 pAspValGluGlyIleAspAlaLeuArgTrpAspAspGlnGluLysIleA 84
300 CGATGTTGAAGGGATAGATGCACTTAGATGGGATGATCAAGAGAAGATAC 349
84 rgAsnTyrValGlySerAlaSerAlaGlyThrSerSerThrAlaAlaPro 100
350 GAAACTACCTGGGAGTGCCTCAGCTGGTACAAAGTTCACAGCTGCTCCT 399
101 ProGluLysCysThrIleGluIleAlaProSerAlaArgThrSerCysAr 117
400 CCTGAAATGTACAAATTGAGATTGCTCCATCTGCCCGCTACTTCATGTAG 449
117 gArgCysSerGluLysIleThrLysGlySerValArgLeuSerAlaLysL 134
450 ACGATGCAAGTGAAGAGATTACAAAAGGATCGGTCCGCTCTTCAGCTAAC 499
134 euGluSerGluGlyProLysGlyIleProTyrHisAlaAsnCysPhe 150
500 TTGAGAGTGAAGTCCCAAGGGTATACATGGTATCATGCCACTGTTTC 549
151 PheGluValSerProSerAlaThrValGluLysPheSerGlyTrpAspTh 167
550 TTTGAGGATATCCCGCTCTGCAACTTTGAGAAGTTCACAGCTGGGATAC 599
167 rLeuSerAspGluAspLysArgThrMetLeuAspLeuValLysLysaspv 184
600 TTTGCCATGAGGATAGAGAACCATGCTCGATCTTGTAAAAAAGATG 649
184 alGlyAsnAsnGluGlnAsnLysGlySerLysArgLysLysSerGluAsn 200
650 TTGGCAACAATGAACAAAATAAGGTTCCAGCCCAAGAAAGTGAAT 699
201 AspIleAspSerTyrLysSerAlaArgLeuAspGluSerThrSerGluG 217
700 GATATTGATGCTACAAATCCGCCAGGTAGATGAAAGTACATCTGAAG 749
217 yThrValArgAsnLysGlyGlnLeuValAspProArgGlySerAsnThrS 234
750 TACAGTCCGAACCAAGGGCAACTTTAGACCCACGTTGGTTCCAAATACTA 799
234 erSerAlaAspIleGlnLeuLysLeuLysGluGlnSerAspThrLeuTrp 250
800 GTTCAGCTGATATCAACTAAAGCTTAAGGAGCAAAAGTGACACACTTTGG 849
251 LysLeuLysAspGlyLeuLysThrHisValSerAlaAlaGluLeuArgAs 267
850 AAGTTAAAGGATGACATTAGACTCATGTATCGGCTGCTGAATTAAGGGA 899
267 pMetLeuGluAlaAsnGlyGlnAspThrSerGlyProGluArgHisLeuL 284
900 TATCTCTGAGCTTAATGGCAGGATACATCAGGACCAGAAAGGCACCTAT 949
284 euAspArgCysAlaAspGlyMetLeuPheGlyAlaLeuGlyProCysPro 300
950 TGGATCGCTGGCGGATGAATGCTATTTGGAGCGCTGGGTCTCTGCCCA 999

301 ValCysAlaAsnGlyMetTyrTyrTyrAsnGlyGlnTyrGlnCysSerG 317
1000 GTCTGTGCTAATGCGATGCTACTATTATAAGTGCAGTACCAATCAGTGG 1049
317 yAsnValSerGluTrpSerLysCysThrTyrSerAlaThrGluProValA 334
1050 TAATGTGTGTCAGAGTGGTCCAAAGTGTACATCTCTGCCACAGAACCTGTC 1099
334 rgValLysLysTrpGlnIleProHisGlyThrLysAsnAspTyrLeu 350
1100 GCGTTAAAGAAAGTGGCAAAATCCACATGACAAAGAAATGATTACCTT 1149
351 MetLysTrpPheLysSerGlnLysValLysLysProGluArgValLeuPr 367
1150 ATGAAGTGGTTCAAATCTCAAAAGGTTAAGAAACCCAGAGGGGTTCTTCC 1199
367 oProMetSerProGluLysSerGlySerLysAlaThrGlnArgThrSerL 384
1200 ACCAATGTCACCTGAGAAATCTGGAAGTAAAGCAACTCAGAGAACATCAT 1249
384 euLeuSerSerLysGlyLeuAspLysLeuArgPheSerValValGlyGln 400
1250 TGCTGTCTTCTAAAGGGTGGATAAAATTAAGTGTTCCTGTGTAGGACAA 1299
401 SerLysGluAlaAlaAsnGluTyrIleGluLysLeuLysLeuAlaGlyAl 417
1300 TCAAAAGAACACCAATGAGTGGATGAGAAGCTCAAACTTCTGTGTGTC 1349
417 aAsnPheTyrAlaArgValValLysAspIleAspCysLeuIleAlaCysG 434
1350 CAACCTCTATGCCAGGTTGTCAAAGATATTGATTGTTTAATTGTCATGTG 1399
434 lyGluLeuAspAsnGluAsnAlaGluValArgLysAlaArgArgLeuLys 450
1400 GTGAGCTCGACAATGAAATGCTCAAGTCCAGGAAGCAAGAGGCTGAAG 1449
451 IleProIleValArgGluGlyTyrIleGlyGluCysValLysArgThrL 467
1450 ATACCAATGTAAGGAGGGTTCATCTGGAGATGTGTTAAAAAGAACAA 1499
467 yScCysHisLeuIleCysIleAsn..TrpAsnAlaLeuGluSerSerL 483
1500 AATGCTGCCATTTGATTGTATAAACTAGAGAATGCTTAGAGTCTCTCAA 1549
483 yGlyLysThrValThrValLysValLysGlyArgSerAlaCysSer*** 499
1550 AAGGCAGTACTGTCACTGTATAAGTTAAGGGCCGAAAGTCTCTTCATGAG 1599
500 SerPro***ValCysLysAsnThrAlaHisIlePro****TrpGluL 516
1600 TCCTCTGGTTTG...CAAGATACGTCTCACATCTCTGA..AGATGGGAAA 1645
516 sHisIleGlnCys***LeuLysHisValLeuThr***His***ValCys. 532
1646 GCATATACAAATCAACCTTAAACATGCTCGACCTGCG...ACTAGGTGTG 1692
533 ThrGlyTyrTyrValLeuGlnIleIleGluGlnAspAspGlySerGluCy 549
1693 AACAGCTACTATGTACTCCAGATCATTAACAGGATGATGGGTCTGAGTG 1742
549 sTyrValPheArgLysTrpGlyArgValGlySerGluLysIleGlyGly 566
1743 CTACGTATTTCGTAAGTGGGACGGGTTGGGAGTGCAGAAAATTGGAGGCG 1792
566 InLysLeuGluGluMetSerLysThrGluAlaIleLysGluPheLysArg 582
1793 AAAAAGTGGAGAGATGCAAAACTGAGGCAATCAAGGAATTTCAAAAGA 1842
583 LeuPheLeuGluLysThrGlyAsnSerTrpGluAlaTrpGluCysLysTh 599
1843 TTATTCTTGAGAACACTGGAAACTCATGGGAGCTTGGGAATGTAAAC 1892

```

599 rAsnPheArgLysGlnProC
1893 CAATTTTCGGAAGCAGCCT
616 lyValLysLysAlaProLy
1943 GTCTTAAGAAAGCACCANJ
633 LeuAlaProGlnLeuLeu
1993 CTTGCTCCTCAATTGCTA
649 rTyrArgAlaAlaMetGluPheGlu***
2043 ATATAGAGCTGCTATGATGAATTTGAAATTA
666 euGlyLysLeuSer*****AsnIleGlu***Gly...eGluAlaLeuThr
2093 TTGGGAAGCTAAGCAAGAAATATTGAGAAAGGATTTGAAGCATTAACCT 2142
683 *****LeuPheGluGlyHisArg***SerSerThrGlyLeu** 699
2143 GAGATACAGAAATTTATT.GAAGGACACCGCTGATCAAGCAGCTGGCTG.TT 2190
699 *GluLysAla**LeuLeu*****PheSerLeuLeuSerL 716
2191 AGAGAAAGCTTAATTGTTGCTGCGAGCAATCGCTTTTTCACCTTTATCCC 2240
716 euLeuPheIleLeuLeuLeuTyrGlyMetArgMetIleSerTyrSerLys 732
2241 TTCTATTATCATCTCATATTATACGGGATGAGGATGATTG.ATGATCAAA 2289
733 AlAlysMetLeuGluAlaLeuGlnAspIleGluIleAlaSerLysIleVa 749
2290 GCGAAAATGCTTGAAGCTCTGCGAGGATTAATGAAATTCCTTCAAGATAGT 2339
749 lGlyPheAspSerAspSerAspGluSerLeuAspAspLysTyrMetIylsL 766
2340 TGGCTTCGATAGCAGCAGTATGATGAATCTCTGATGATAAATATATAAAC 2389
766 euHisCysAspIleThrProLeuAlaHisaspSerGluAspTyrLysLeu 782
2390 TTCACTGTGACATCACCCGCTGGCTCAGCATAGTGAAGATTAACAAGTTA 2439
783 lIleGluGlnTyrLeuLeuAsnThrHisAlaProThrHisLysAspTrpSe 799
2440 ATTGAGCAGTATCTCCCAACACACATGCTCTACTCACAGGACTGGTC 2489
799 rLeuGluLeuGluValPheSerLeuAspArgAspGlyGluLeuAsnL 816
2490 GCTGGAAGCTGGAGGAAGTTTTTTCACCTGATCGAGATGGAGAACTTAATA 2539
816 ysTyrSerArgTyrLysAsnAsnLeuHisAsnLysMetLeuLeuTyrPheHis 832
2540 AGTACTCAAGATATAAAAAATTAATCGCATACAAGATGCTATTATGGCAC 2589
833 GlySerArgLeuThrAsnPheValGlyLysLeuSerGluGlyLeuArgI 849
2590 GGTCAAGGTTGACGAATTTTGGGGAATCTTCTAGTCAAGGCTAAGAAT 2639
849 eAlaProGluAlaProValThrGlyTyrMetPheGlyLysGlyLeuT 866
2640 TGCACCTCTCAGGCACCTGTTTACTGGCTATATGTCGCAAGGCTCT 2689
866 yrPheAlaAspLeuValSerLysSerAlaGlnTyrCysTyrValAspArg 882
2690 ACTTTCAGATCTAGTAAGCAAGACCGCACAAATCTGTTATGCGATAGG 2739
883 AsnAsnProValGlyLeuMetLeuLeuSerGluValAlaLeuGlyAspMe 899
2740 ATAATCTCTGAGGTTGATGCTCTTCTGAGGTTGCTTTAGGAGACAT 2789
899 tTyrGluLeuLysLysAlaThrSerMetAspLysProArgGlyLysH 916

```

```

2790 GTATGAACTAAAGAAAGCCACGTCATGGACAAACCTCCAAGAGGGAAGC 2839
916 isSerThrLysGlyLeuGlyLysThrValProLeuGluSerGluPheVal 932
2840 ATTCGACCAAGGATTAGGCAAAACCGTGGACCTGGAGTTCAGATTGTG 2889
933 LysTyrArgAspAspValValProCysGlyLysProValProSerSe 949
2890 AAGTGGAGGATGATGTCGTAGTTCCCTGCGCAAGCGGTGCCATCATC 2939
949 rIleArgSerSerGluLeuMetTyrAsnGluTyrIleValTyrAsnThrS 966
2940 AATTAGGAGCTCTGAACATCATGTACAAATGAGTACATCTGTACACACAT 2989
966 erGlnValLysMetGlnPheLeuLeuLysValArgPheHisHisLysArg 982
2990 CCCAGGTGAAGATGAGTCTTCTGCTGAAGTTCGCTTCCATCACAAAGAGG 3039
seq_name: gb_pl:ZMPARP2
seq_documentation_block:
LOCUS ZMPARP2 3211 bp mRNA PLN 19-NOV-1997
DEFINITION Zea mays mRNA for poly(ADP-ribose) polymerase (3211bp).
ACCESSION AJ222589
VERSION AJ222589.1 GI:2632128
KEYWORDS PARP gene; poly(ADP-ribose) polymerase.
SOURCE Zea mays.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PAC
clade; Panicoidae; Andropogoneae; Zea.
REFERENCE
AUTHORS Bahiychuk, E., Cottrell, P., Storozhenko, S., Fuangthong, M.,
O'Farrell, M., Van Montagu, M., Inze, D. and Kushnir, S.
TITLE Higher plants possess two poly(ADP-ribose) polymerases
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3211)
AUTHORS Kushnir, S.
TITLE Direct Submission
JOURNAL Submitted (14-NOV-1997) Kushnir S., VIB, Dep.Genetics,
Ledeganckstraat 35, Gent, B9000, Belgium
FEATURES
source
1..3211
/organism="Zea mays"
/db_xref="taxon:4577"
113..3022
/gene="PARP"
113..3022
/gene="PARP"
/function="secondary protein modification"
/codon_start=1
/product="poly(ADP-ribose) polymerase"
/protein_id="CAA10889.1"
/db_xref="GI:2632128"
/db_xref="SPTREMBL:O24570"
/translation="MAAPPAKWADEYAKSGRASCSCRSPIAKDOLGLKQWASQFD
GMPWNHASVDDVEGIDALURKADQEKIRNVSASAGTSTAAPPKCIEIETAPSAR
TSCRCSEKITKGSVRLSAKLESEGPKGIPWYHANCFFVSPSATVEKFSQWDTLSD
DKRTMLDLVKDQVGNQKSKRKGKSFNDIDYSKARLDESTSEGTVRNKGQLVDR
GSNTSSADIQILKEQSDTLWKLDGLKTHVSAALRDLMEANGODTSGPRLHLLRC
ADGMIFGALGCPVCANGVYNGYOCSGNVSEWSTYTSATPEPVAVKKWQIPIHGT
ANDYLMKFRSKQKVKKPERVLPPMSPEKSKATQRTSLSSKGLDKLRFVSVQSK
AANEMIEKLKLAGANFYARVVVDIDCLACELNENAEVRKARRLKIPVREGIGE
CVKNKMLPFPLYKLENALESKSTVTYVKVKSNAVESSEGLQDTAHILEDKSYIN
ATLNMSDALGNSWYVLIITEODDGGSEYVFRKWRGVGSEKIGQKLEMSKTEAIK
EFKRLFEKTKNSWEAWCKTFRKQRFYPLDQVGVKAPKRDISEMKSLSAPQ
LLELMKMLFNVTYRAAMEFEINNSEMPGLKSKENIEKFEALTEIQNLKIDTADQ
ALAVRESLIVAASNRFTLIPS.IHPHIIIRDEDDLMIKAKMLEALQDIEIAKIVGFS
DSDESDDKYMKLHCDITPLAHDSEDKLIEQYLLNTHAPTHDKMSLELEVESLDRD
GELNKSRYKNNHKNMLLWHSRLTNFVGLISQGLRLAPPEAPVPTGYMGKGLYFAD
LVKSAQYCYVDNRNPNVGLMLLSVALGMDKATSMKPPKPKHSTKGLGKTVP
ESEFVKWRDDVVVPCGKVPVSSIRSELNNEYIVINTSQVKMQLKVRPHHKK"

```

BASE COUNT 968 a 604 c 813 g 826 t
 ORIGIN
 alignment_scores:
 Quality: 4666.50 Length: 985
 Ratio: 4.980 Gaps: 6
 Percent Similarity: 95.127 Percent Identity: 94.416
 alignment_block:
 US-09-236-995D-2 x ZMPARP2 ..
 Align seg 1/1 to: ZMPARP2 from: 1 to: 3211
 1 MetAlaAlaProProLysAlaTrpLysAlaGluTyrAlaLysSerGlyAr 17
 113 ATGGCGGCGCGCCCAAGGCGTGGAGGCGGAGTATGCCAAGTCTGGGCG 162
 17 gAlaSerCysLysSerCysArgSerProLeAlaLysAspGlnLeuArgL 34
 163 GGCGCTCGTCAAGTCATGCGCGTCCCTATCGCCAAAGGACCAGCTCCGTC 212
 34 euGlyLysMetValGlnAlaSerGlnPheAspGlyPheMetProMetTrp 50
 213 TTGGCAAGATGGTTTCAGGCGTCACAGTTTCGACGGCTTCATGCGGATGG 262
 51 AsnHisAlaArgCysIlePheSerLysLysAsnGlnIleLysSerValAs 67
 263 AACCATGCC.....AGCGTTGA 279
 67 pAspValGluGlyIleAspAlaLeuArgTrpAspGlnGluLysIleA 84
 280 CGATGTTGAAGGATAGATGCACCTTAGATGGGATGATCAAGAGAAGATAC 329
 84 rGAsnTyrValGlySerAlaSerAlaGlyThrSerThrAlaAlaPro 100
 330 GAACACTACGTTGGAGTGCTCAGCTGGTACAGATTCTACAGCTGCTCCT 379
 101 ProGluLysCysThrIleGluIleAlaProSerAlaArgThrSerCysAr 117
 380 CCTGAGAAATGTACAATTGAGATTGCTCCATCTGCCCGTACTTTCATGTAG 429
 117 qArgCysSerGluLysIleThrLysGlySerValArgLeuSerAlaLysL 134
 430 ACGATGACGTGAAAGATTACAAAAAGGATCGGTCGCTCTTCAGCTAAGC 479
 134 euGluSerGluGlyProLysGlyIleProTrpTyrHisAlaAsnCysPhe 150
 480 TTGAGATGAAGGTCCCAAGGGTATACCATGGTATCATGCCAAGCTGTTTC 529
 151 PheGluValSerProSerAlaThrValGluLysPheSerGlyTrpAspTh 167
 530 TTTGAGGTATCCCGTCTGCAACTGTTGAGAAGTTCTCAGGCTGGGATAC 579
 167 rLeuSerAspGluAspLysArgThrMetLeuAspLeuValLysLysAspV 184
 580 TTTGTCCGATGAGGATAAGAACCATGCTCGATCTTGTAAAAAAGATG 629
 184 aGlyAsnAsnGluGlnAsnLysGlySerLysArgLysLysSerGluAsn 200
 630 TTGGCAACAATGAACAAAAATAAGGTTTCCAGCGCAAGAAAGTGAAT 679
 201 AspIleAspSerTyrLysSerAlaArgLeuAspGluSerThrSerGluG 217
 680 GATATTGATAGCTACAAATCCGCCAGGTAGATGAAAGTACATCTGAAGG 729
 217 yThrValArgAsnLysGlyGlnLeuValAspProArgGlySerAsnThrS 234
 730 TACAGTGGCAACAAAGGCAACTTGTAGACCCACGCTGGTTCCAATACTA 779
 234 erSerAlaAspIleGlnLeuLysLysGluGlnSerAspThrLeuTrp 250
 780 GTTCAGCTGATATCCAACATAAGCTTAAGGAGCAAAAGTGACACACTTTGG 829

251 LysLeuLysAspGlyLeuLysThrHisValSerAlaAlaGluLeuArgAs 267
 830 AAGTTAAAGGATGGACTTAAGACTCATGTATCGGCTGCTGAATTAAGGA 879
 267 pMetLeuGluAlaAsnGlyGlnAspThrSerGlyProGluArgHisLeuL 284
 880 TATGCTTGAGGCTAATGGCAGGATACATCAGGACCAGAAAGGACCTAT 929
 284 euAspArgCysAlaAspGlyMetLeuPheGlyAlaLeuGlyProCysPro 300
 930 TGGATCGCTGCGGATGGAATGATATTGAGCGCTGGTCTCTTGCCCA 979
 301 ValCysAlaAsnGlyMetTyrTyrAsnGlyGlnTyrGlnCysSerG 317
 980 GTCTGTGCTAATGGCATGTACTATTATTAATGGTCAGTACCAATGACGTG 1029
 317 yAsnValSerGluTrpSerLysCysThrTyrSerAlaThrGluProValA 334
 1030 TAATGTGTACAGTGGTCCAAGTGTACATCTGTGCCACAGAACCTGTCC 1079
 334 rgValLysLysLysTrpGlnIleProHisGlyThrLysAsnAspTyrLeu 350
 1080 GCGTTAAGAGAAGTGGCAAAATTCACATGCAACAAAGAATGATTACCTT 1129
 351 MetLysTrpPheLysSerGlnLysValLysLysProGluArgValLeuPr 367
 1130 ATGAAGTGGTTCAAAATCTCAAAAGGTTAAGAAACCCAGAGGGTCTTCC 1179
 367 oProMetSerProGluLysSerGlySerLysAlaThrGlnArgThrSerL 384
 1180 ACCAATGTCACTGAGAAATCTGGAAGTAAAGCAACTCAGAGAACATCAT 1229
 384 euLeuSerSerLysGlyLeuAspLysLeuArgPheSerValValGlyGln 400
 1230 TGCTGTCTTCTAAAGGGTTGGATAAAATTAAGGTTTCTGTGTAGGACAA 1279
 401 SerLysGluAlaAlaAsnGluTrpIleGluLysLeuLysLeuAlaGlyAl 417
 1280 TCAAAAGAACGACCAATGAGTGGATTGAGAAGCTCAAACTTCTGTGTGC 1329
 417 aAsnPheTyrAlaArgValValLysAspIleAspCysLeuIleAlaCysG 434
 1330 CAACTTCTATGCCAGGGTTGTCAAAGATATTGATTTGTTAAATGCATGTG 1379
 434 lyGluLeuAspAsnGluAsnAlaGluValArgLysAlaArgArgLeuLys 450
 1380 GTGAGCTCGACATGAAATGCTGAGTCAAGGAAAGCAAGAGGCTGAAG 1429
 451 IleProIleValArgGluGlyTyrIleGlyGluCysValLysArgThrL 467
 1430 ATACCAATTTAAGGGAGGGTTACATTGGAGAATGTGTTAAAGAAACAA 1479
 467 yCysCysHisLeuIleCysIleAsn..TrpAsnAlaLeuGluSerSerL 483
 1480 AATGCTGCCATTTGATTTGATATAAAGTATAGAAATGCTTAGAGTCTCAA 1529
 483 ysGly***ThrValThrValLysValLysGlyArgSerAlaCysSer*** 499
 1530 AAGCGAGTACTGTCATGTTAAAGTTAAGGCGCAAGTGTCTGTTCAAGAG 1579
 500 SerPro***ValCysLysAsnThrAlaHisIlePro****TrpGluLys 516
 1580 TCCTCTGGTGTG...CAAGATACCTGCTACATTTCTTGA..AGATGGGAAA 1625
 516 sHisIleGlnCys***LeuLysHisValLeuThr***His***ValCys. 532
 1626 GCATATACAAATGCACTTAAACATGCTGACCTGGC...ACTAGGTGTG 1672
 533 ThrGlyTyrTyrValLeuGlnIleIleGluGlnAspGlySerGluCys 549
 1673 AACAGCTACTATGCTACCCAGATCATTTGACAGGATGATGGGTCTGAGTG 1722

549 sTyrValPheArgLysTrpGlyArgValGlySerGluLysIleGlyGly 566
|||||
1723 CTACGTATTTCTAGTGGGACGGTGGGAGTCAGAAAATTCGAGGGC 1772
|||||
566 InLysLeuGluCLeuMetSerLysThrGluAlaIleLysGluPheLysArg 582
|||||
1773 AAAAAGCTGGAGGAGATGTCAGAAATCGAGCAATCAAGGAATCAAAAGA 1822
|||||
583 LeuPheLeuGluLysThrGlyAsnSerTrpGluAlaTrpGluCysLysTh 599
|||||
1823 TTATTTCTTGAGAGACTGGAACTCATGGGAAGCTGGGAATGTAAAC 1872
|||||
599 rAsnPheArgLysGlnProGlyArgPheTyrProLeuAspValAspTyrG 616
|||||
1873 CAATTTTCGGAAGCAGCCCTGGGAGATTTTACCACCTTCATGTTGATTG 1922
|||||
616 yValLysLysAlaProLysArgLysAspIleSerGluMetLysSerSer 632
|||||
1923 GTGTTTAAAGAAAGCACCAAGCGGAAGATATCAGTGAATGAAAAGTTCT 1972
|||||
633 LeuAlaProGlnLeuLeuGluMetLysMetLeuPheAsnValGluTh 649
|||||
1973 CTTCCTCTCAATTCCTAGAACTCATGAGATGCTTTTCAATGTGGAC 2022
|||||
649 rTyrArgAlaLaMetMetGluPheGlu***AsnMetSerGluMetProL 666
|||||
2023 ATATAGAGCTGCTATGATGGAATTTGAAATTAATATATGTCAGAAATGCCTC 2072
|||||
666 euGlyLysLeuSer*****AsnIleGlu***GlyPheGluAlaLeuThr 682
|||||
2073 TTGGGAAGCTAAGCAAGGAAATATTTGAGAAGGATTTGAAGCATTAAC 2122
|||||
683 *****LeuPheGluGlyHisArg***SerSerThrGlyLeu** 699
|||||
2123 GAGTACAGAAATTTATT.GAGGACACCGCTGATCAAGCATCGCTG.TT 2170
|||||
699 *gluLysAla***LeuLeu*****PheSerLeuLeuSerL 716
|||||
2171 AGAGAAAGCTTAATTTGTTGCTGCGAGCAATCGCTTTTTCACCTTATCCC 2220
|||||
716 euLeuPheIleLeuIleLeuTyrGlyMetArgMetIleSerTyrSerLys 732
|||||
2221 TTCATATTATCTCATATTATACGGGATGAGGATGATTG.ATGATCAA 2269
|||||
733 AlaLysMetLeuGluAlaLeuGlnAspIleGluIleAlaSerLysIleVa 749
|||||
2270 GCGAAATGCTTGAAGCTCTCGAGATATTGAAATTTGCTTCAAGATAGT 2319
|||||
749 lcyPheAspSerAspSerAspGluSerLeuAspLysTyrMetLysL 766
|||||
2320 TGGCTTCGATAGCGACAGTGAATCTTTGATGATATAATATATGAAC 2369
|||||
766 euHisCysAspIleThrProLeuAlaHisAspSerGluAspTyrLysLeu 782
|||||
2370 TTCATCTGACATACCCCGCTGGCTACGATAGTGAAGATTACAAGTTA 2419
|||||
783 ileGluGlnTyrLeuLeuAsnThrHisAlaProThrHisLysAspTrpSe 799
|||||
2420 ATTGAGCAGTATCTCTCAACACACATGCTCTACTCACAGGACTGGTC 2469
|||||
799 rLeuLeuLeuGluAlaValPheSerLeuAspArgAspGlyGluLeuAsnL 816
|||||
2470 GCTGGAAGCTGGAGGAAGTTTTCCTACTTGATCGAGATGGAGAACTTAATA 2519
|||||
816 yTyrSerArgTyrLysAsnAsnLeuHisAsnLysMetLeuLeuTrpHis 832
|||||
2520 AGTACTCAAGATATAAAATATTCGATATAACAAGATGCTATTATGGCAC 2569
|||||
833 GlySerArgLeuThrAsnPheValGlyIleLeuSerGlnGlyLeuArgIl 849
|||||
2570 GGTTCAGGTTGACGAAATTTTGTGGGAATTTCTTAGTCAAGGCGCTAAGAT 2619
|||||
849 eAlaProProGluAlaProValThrGlyTyrMetPheGlyLysGlyLeu 866

|||||
2620 TGACACCTCTGAGGACCTGTTACTGGCTATATGTTCCGCAAGGCCTCT 2669
|||||
866 yRheAlaAspLeuValSerLysSerAlaGlnTyrCysTrValAspArg 882
|||||
2670 ACTTTGGAGATCTAGTAAGCAAGAGCGCACATACTGTTATGTGGATAGG 2719
|||||
883 AsnAsnProValGlyLeuMetLeuLeuSerGluValAlaLeuGlyAspMe 899
|||||
2720 AATAATCTCTAGGTTTGATGCTTCTTCTGAGGTTGCTTTAGGAGACAT 2769
|||||
899 tTyrGluLeuLysLysAlaThrSerMetAspLysProProArgGlyLysH 916
|||||
2770 GTATGAACCTAAAGAAAGCCACGTCATGGCAAAACCTCCAAGAGGAAAGC 2819
|||||
916 isSerThrLysGlyLeuGlyLysThrValProLeuGluSerGluPheVal 932
|||||
2820 ATTTCGACCAAGGATTAGGCAAAACCGTCCACTGGAGTCAGAGTTTGTG 2869
|||||
933 LysTrpArgAspValValValProCysGlyLysProValProSerSe 949
|||||
2870 AAGTGGAGGATGATGTCGTAGTTCCCTGCGCAAGCGGTGCATCATC 2919
|||||
949 rIleArgSerSerGluLeuMetTyrAsnGluTyrIleValTyrAsnThrS 966
|||||
2920 AATTAGAGCTCTCAACTCATGTACAAATGAGTACATCGTCTACACACAT 2969
|||||
966 erGlnValLysMetGlnPheLeuLysValArgPheHisLysLysArg 982
|||||
2970 CCAGGTGAAGATGCGAGTTCTTCTGAAGTGCCTTCCATCATCAAGAGG 3019

seq_name: gb_pl:ATH131705

seq_documentation_block:

LOCUS ATH131705 3187 bp mRNA PLN 18-DEC-1998
DEFINITION Arabidopsis thaliana mRNA for poly(ADP-ribose) polymerase.

ACCESSION AJ131705

VERSION AJ131705.1 GI:4038490

KEYWORDS NAD(+) ADP-ribosyltransferase; parp-1 gene; poly(ADP-ribose) polymerase.

SOURCE thale cress.

ORGANISM

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE 1 (bases 1 to 3187)

AUTHORS Kazmaier,M.

TITLE Direct Submission

JOURNAL

Submitted (17-DEC-1998) Kazmaier M., Departement D'Ecophysiologie Vegetale Et De Microbiologie, Commissariat A L'energie Atomique, CEA Cadarache, Lab. De Radiobiol. Vegetale, Bat. 185, 13108 St. Paul-les-Durance, FRANCE

REFERENCE 2 (bases 1 to 3187)

AUTHORS Doucet-chabeaud,G. and Kazmaier,M.

JOURNAL

FEATURES

source

Location/Qualifiers

1..3187

/organism="Arabidopsis thaliana"

/cultivar="landsberg erecta"

/db_xref="taxon:3702"

11..2963

/gene="parp-1"

11..2963

/gene="parp-1"

/standard_name="NAD(+) ADP-ribosyltransferase"

/EC_number="2.4.2.30"

/function="ADP-Ribose polymer synthesis"

/note="mol. weight: 111 kDa"

/codon_start=1

/evidence=experimental

/product="poly(ADP-ribose) polymerase"

/protein_id="CAA10482.1"

/db_xref="GI:4038491"

1313 ACAAGCTGTTGGTGTGTCGACATGCAGCATATACAGACGCCGTGAAT 1362

443 largLysAlaArgLeuLysIleProIleValArgGLuglyTyrlleg 460
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1363 CAGAAAGCAAGGAGGATGAACTGGCAATCGTGAGACGAGGATTATTTGG 1412

460 lyGLuCysValLysArgThyLysCysCysHisLeulleCysIleAsnTrp 476
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1413 TTGATTTGTTTAAAAACAAGAGG.....AAACTTCCATTTGCACAAG 1453

477 AsnAlaLeuGlusSerLysGly***ThrValThrValLysValLysGI 493
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1454 TACAAAATTGAAGACACTAGTGAGAGCCCTGTCACTGTTAAAGTAAGAAG 1503

493 yArGSerAlaCySSer***SerPro***ValCysLysAsnThrAlaHisi 510
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1504 ACGAAGCGCT.....GTGCATG 1520

510 lePro*****TrpGLuLysHisIleGlnCys***Leu..... 522
||| :|||:||||| ||| :|||
1521 AAGCGCTGGGCTCCAAGAGCAC.....TGTCACATCTTGAAGATGGG 1564

523LysHisValLeuThr***His**ValCystThrCly.. 534
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1565 AACAGTATCTATAACAACAACCTCTGAGCATCTCTGATCTCTACCGGTAT 1614

535TyrrValLeuGlnIleIleGluInAspAspGlySerGLuc 549
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1615 CAATAGTTATTACATACATCCAGATAATCAAGAAGATAAAGTTAGATT 1664

549 ysTyrrValPheArgLysTrpDlyArgValGlySerGLuylsleGlyGly 565
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1665 GTTAGCTATTTCGTAATGGGCCGCGCTTGGAAATGAAAAGATTGGTGST 1714

566 GlnLysLeuGLuMetSerLysThrGLuAlaIleLysGLuPhelysar 582
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1715 AACAAAGTGGAGGAATGTCAAAGCTGTGCGGGTTCACGAATCAAACG 1764

582 gLeuPheLeuGLuLysThrGlyYasnSerTrpGLuAlaTrpGLuCysLysT 599
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1765 TCTATTCTTGAANAACC CGAACACATGGGAATCTTGGGACACAAAAA 1814

599 hrAsnPheArgLysGlnProGlyArgPheTyrrProLeuAspValAspTyrr 615
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1815 CGAATTTTCAGAAACACACTCGAAATTTCCCGTTTGGACATTGATTAT 1864

616 GlyValLysLysAlaProLysArgLysAspIleSerGLuMetLysSerSe 632
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1865 GGAGTTAATACCAAGTAGCCCAAAAAGAGCCATTTCAGACCAGTAGCAA 1914

632 rLeuAlaProGlnLeuLeuGLuLeuMetLysMetLeuPheAsnValGluTr 649
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1915 CCTTGCTCCATCATTAATAGAAATTGATGAAGATGCTTTTTTGATGTGGAA 1964

649 hrTyrrArgAlaAlaMetMetGLuPheGLu***AsnMetSerGLuMetPro 665
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1965 CTTACAGATCTCGAATGATGGAGTTTCGAGATAAATATGTCAGAGATGCCA 2014

666 LeuGLyLysLeuSer*****AsnIleGlu***GlyPheGLuAlaLeuTh 682
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2015 CTTGGGAAGCTCAGCAAAACATAATATACAGAAGGGTTTTCAGGCATTGAC 2064

682 r*****LeuPheGLuGLyHis.Arg***SerSerThrGlyLeu 698
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2065 GGAGATACAGAGCTATTG.....ACTGAAGGCGACCCCCAGCGCTACTA 2108

699 **GLyLysAla***LeuLeu*****PeSerLeuLeuSe 715
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2109 TGAAGAAAGACTTGCTGTGTGATGCTAGTAACAGATTTTACCATGATC 2158

715 rLeuLeuPheIleLeuLeuTyrrGlyMetArgMetIleSerTyrrSerL 732
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2159 CCTTATTATTCCTCATATATCCGAGATGAAGATCACTTT.AAGTCAA 2207

[illegible]

seq name: qb_pl:AC006593

seq_documentation_block:

seq_documentation_block,		PIN	05-APR-2000
LOCUS	AC006593	79663 bp	DNA
DEFINITION	Arabidopsis thaliana chromosome II section 176 of 255 of the complete sequence. Sequence from clones F16D14, T28P16.		
ACCESSION	AC006593 AE002093		
VERSION	AC006593.4	GI:6598616	

KEYWORDS
SOURCE
ORGANISM

HTG.
thale cress.
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 79663)

REFERENCE
AUTHORS

Lin, X., Kaul, S., Rounsley, S.D., Shea, T.P., Benito, M.-I., Town, C.D.,
Fuji, C.Y., Mason, T.M., Bowman, C.L., Barnstead, M.E.,
Felchlyum, T.V., Buell, C.R., Ketchum, K.A., Lee, J.J., Ronning, C.M.,
Koo, H., Moffat, K.S., Cronin, L.A., Shen, M., VanAken, S.E., Unayam, L.,
Tallon, L.J., Gill, J.E., Adams, M.D., Carrera, A.J., Creasy, T.H.,
Goodman, H.M., Somerville, C.R., Copenhaver, G.P., Preuss, D.,
Nierman, W.C., White, O., Eisen, J.A., Salzberg, S.L., Fraser, C.M. and
Venter, J.C.

TITLE

Sequence and analysis of chromosome 2 of the plant Arabidopsis

JOURNAL
MEDLINE
PUBMED
REFERENCE

Nature 402 (6763), 761-768 (1999)

20083487

10617197

2 (bases 1 to 79663)

Lin, X.

Direct Submission

JOURNAL
COMMENT

Submitted (09-Mar-2000) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA
On Dec 17, 1999 this sequence version replaced gi:4432811.
The sequence and annotation of chromosome 2 were merged from those
of the individual clones on this chromosome after removing
overlaps. For detailed information, please see the TIGR web site
(<http://www.tigr.org/tdbat/at.html>).

Genes were identified by a combination of three methods: Gene
prediction programs including GRAIL
(<ftp://arthur.epm.ornl.gov/pub/xgrail>), GeneFinder (Phil Green,
University of Washington), Genscan (Chris Burge,
<http://genomic.stanford.edu/genscan.html>), and NetPlantGene
(<http://www.cbs.dtu.dk/services/NetGene2/>), searches of the
complete sequence against a peptide database and plant EST
databases at TIGR, and manual curations based on those analyses.
Annotated genes are named to indicate the level of evidence for
their annotation. Genes with similarity to other proteins are named
after the database hits. Genes without significant peptide
similarity but with EST similarity are named as 'unknown' proteins.
Genes without protein or EST similarity, that are predicted by two
or more gene prediction programs over most of their length are
annotated as 'hypothetical' proteins. Genes encoding tRNAs are
predicted by tRNAscan-SE (Sean Eddy,
<http://genome.wustl.edu/eddy/tRNAscan-SE/>). Simple repeats were
identified by RepeatMasker (Arian Smit,
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>). Genes are
numbered from the top to bottom of the chromosome.

We thank the CSHL/WashU/ABI consortium for sequencing BAC clones
F6P23, F5J6, T17A5, and T13L16, the ESSA group for sequencing clone
F13D4, and Scott Jackson, Jiming Jiang, Klaus Meyer, Eric Richards
and Satoshi Tabata for helpful assistance. In addition, we would
like to thank the TIGR Bioinformatics Department, especially Lixin
Zhou, Hanif Khalak, Michael E. Heaney, Lily Fu, Feng Liang, Jeremy
Peterson, Michael Holmes, and Delwood Richardson for software and
database support.

This work was supported by the National Science Foundation,
Department of Energy and the US Department of Agriculture.

Address all correspondence to: at@tigr.org.

FEATURES
source

Location/Qualifiers
1. .79663
/organism="Arabidopsis thaliana"
/cultivar="Columbia"
/db_xref="taxon:3702"
/chromosome="II"
1. .79462
/note="Sequence from clone F16D14"

misc_feature

mRNA

complement(join(<1349..1687,1764..1967,2138..2263,
2342..2623,2709..2879,2963..3124,3898..3961,4063..>4406))
/gene="At2g31170"
complement(<1349..>4406)
/gene="At2g31170"
/note="F16D14.1"

CDS

complement(join(1349..1687,1764..1967,2138..2263,
2342..2623,2709..2879,2963..3124,3898..3961,4063..>4406))
/gene="At2g31170"
/codon_start=1
/product="putative cysteinyl-tRNA synthetase"
/protein_id="AAD20662.1"
/db_xref="GI:4432812"

/translation="MASSVINLFKSCRPETPIRFSSLPKSOFRIOPLRPGKETOLRR
CFITLSTDDGAPISGGKELMHLHNTSKKELFKPKVEGKVMYVCGVTAYDLSHIG
HARVYVTFDLRLYLKHLGYEVSYYRNFDDDKIARAKELEEDPISLRRFCSEFN
RMEQLQCLDPSQVPRVSDHIQIIDLKIOIDONGYAYKVDGDIYFVDKPTCYGLS
GRKLEDNRAGERVAVDTRKKHPADFAALKAKEGPFWESPMGRGPGWHIECSAMSA
AYLGYSFDIHGGMDLVFPHEHEIAOSCAACDSSNIYWHNGPVTVDSEKMSKLG
NEFTTQVIDLHPLALRFLMGTHYRSPINYSDFLLESASERIFYIYOTLHDCESAL
GKGDSTFENGVSFSDTLTSINTFRTEFVMSDDLLTPVTLAAMSEPKTINDLHTTR
VGKQARDEESKALETTIRDLITILGLMPTYSSEVLEQLKEKALRAGLEEDVLQR
VOERTDARKKEYERSDAIRKDLAKVGIALMDSPSGTTRWRPAIPLALQEPVTTTP"
complement(join(5525..6082,6603..6732,6833..>6965))
/gene="At2g31180"
complement(5525..>6965)
/gene="At2g31180"
/note="F16D14.2"
complement(join(5596..6082,6603..6732,6833..>6965))
/gene="At2g31180"
/codon_start=1
/product="putative MYB family transcription factor"
/protein_id="AAD20663.1"
/db_xref="GI:4432813"

repeat_region

/translation="NGRAPCCEKMGVGRGTPWPEDQILINVIHYLHNSNRALPKHA
GILRCGKSCRLRWINYLPRDKRGFTPOEEQITINLHESLGNWSATAALPQRTDN
ETKNVHTHLKRLKSLNNGDQKDVNGINETTNEKGSIVDTASIQOFSNITTF
DISDNKDDIMSYEDISALIDDSFSDSVISYDNSKNKKEKEDGLEIDRNSKKCSYS
NSKLYNDMEFDFVFTSNRRIEFSIDIPEF"
complement(7897..7920)
/rpt_family="POLY_A"
8038..8151
/rpt_family="TAAAAn"
9946..9977

repeat_region

/rpt_family="TAAAn"
complement(join(<10248..10673,10764..10886,10983..11162,
11326..11423,11506..11599,11690..11749,11871..11955,
12057..>12145))
/gene="At2g31190"
complement(<10248..>12145)
/gene="At2g31190"

mRNA

complement(join(10248..10673,10764..10886,10983..11162,
11326..11423,11506..11599,11690..11749,11871..11955,
12057..>12145))
/gene="At2g31190"
complement(<10248..>12145)
/gene="At2g31190"
/note="F16D14.3"

gene

complement(join(10248..10673,10764..10886,10983..11162,
11326..11423,11506..11599,11690..11749,11871..11955,
12057..12145))
/gene="At2g31190"
/note="unknown protein"
/codon_start=1
/protein_id="AAD20664.1"
/db_xref="GI:4432814"

CDS

/translation="MKVVDADARVPQKRVESLKNKFFPGYPYSVNEGYLRYQFRAL
QHSFSAALSVLSTQSLFLAAGLRTPAQATVWSWILKDGQHVKGKLTCSNLGARMDS
PKRWRLADVLDTGLGVSPCLPFLHLEAGLGNFAKGMATVAARATRLPYSF
AKEGNLSIDFAKGEAISTFNAGIGAGIOLASTICSSMEGLVGSILSVHYVSIV
EOMEGVPTINLPORLTIIVANFLKTKGVSPDLPQEDLMPEPIQDAGNVKVR
ALHKAVKPSVQRKLVFVEEKILLSHGSKSDMKVLEHDATGADALRGWVAAYVKSM
TKIYNDDDDIILQDAYDKMNDVFNFLSQVQAQWYTDRLDGTGRFAM"
join(12936..12984,13233..13498,13588..13809)
/gene="At2g31200"
12936..13809
/gene="At2g31200"
/note="F16D14.4"

mRNA

complement(join(12936..12984,13233..13498,13588..13809)
/gene="At2g31200"

gene

complement(join(12936..12984,13233..13498,13588..13809)
/gene="At2g31200"

CDS

complement(join(12936..12984,13233..13498,13588..13809)
/gene="At2g31200"

77318	ACTTCGTTGGAGATCAGCAAAAGATTAGAAAAATATGTCGAATCTCGAG	77266
91	erAlaGlyThrSerSerThrAlaAlaProProGluLysCysThr.....	105
77268	CAGGAGTAACACAAAGCACAAAGCACAGGCACAAAGCACAGCAGTACCGCT	77219
106lleGluIleAlaProSerAlaAlaTh	114
77218	ANTAAATCCCAAACTAGAAATATGGATGTGAAGTGTCCACAACTTCTCTGTC	77169
114	rSerCysArgArgCysSerGluLysIleThrLysGly.....	126
77168	CGSTTGCAGAAAGTGTAGCGAAAGATCTTGAAGAGAGATAAATGTT	77119
126	126
77118	GTTCTACTTCATCTGTCTATGTTGTAGCGTTTATATGAACGTGTTCCGAT	77069
127Ser.ValArgLeuSerA	132
77068	ATCCATTACGTTTGTGTTTTCGATTTTGTTCAGGTACGTAATATCT	77019
132	laLysLeuGluSerGluGlyProLysGlyIleProTptTyrHisAlaAsn	148
77018	CCAAGCCTCAAGGCCCGGTAAACAAAGTTTGATGTGCATCAGCTAA	76969
149	CysPheGluValSerProSerAlaThrValGluLysPheSerGlyTr	165
76968	TGTTCTTGAATGTCTCTCTACTGAACGTGGAAGTTTGTCTGGATG	76919
165	pAspThrLeuSerAspGluAspLysArgThrMetLeuAspLeuValLysI	182
76918	GAGAATATACAGACTCAGACCAAGAGCTCTTCCTCCCTTAACTAGAG	76869
182	ysAspValClyAsnAsnGluGlnAsnLysGlySerLys.....	194
76868	AAGCTCTCGCGCAGCCAAAACCTCGTAAGAGTCTTAAAGTGTACGTAAT	76819
195ArgLysLysSerGluAsnAspIleAspSerTyr.....	205
76818	ATATTGATTCTCTGTACAGGATCCCTGATTTGCAGTACTTTAGTCTGAT	76769
206LysSerAlaArgLeuA	211
76768	CTTTCCTCGTAATTTATTTTGCCCAACTGGGACAGACAGCAGAACAC	76719
211	spGluSerThrSerGluGlyThrValArgAsnLysGlyGlnLeuValAsp	227
76718	GTCAACAAATTAAGAGCAGGCACAAAAGCAAAATGATTCCTGTGAT	76669
228	ProArgGlySerAsnThrSerSerAlaAspIleGlnLeu.....	240
76668	AACGAGAAGTCGAAACTAGCAAAAGTAGTTTGTACATGCTCTACAAGTG	76619
241LysLeuLysGluGlnSerAspT	248
76618	TGCTTTACAACCTTGTAGCAAGAAAGGAATGGAGCCCACTAAAG	76569
248	hrLeuTrpLysLeuLysAspGlyLysThrHisValSerAlaAlaGlu	264
76568	AATTGTGGACCTGAAGGATGATCTCAAAAATATGTAACATCAGCTGAG	76519
265	LeuArgAspMetLeuGluAlaAsnGlyGlnAspThrSerGlyProGluAr	281
76518	TTGCGGAATGCTTGAAGTAATGAACAAGATACAAGAGGATCTGAAC	76469
281	gHisLeuLeuAspArg.....	286
76468	TGATCTCGGTGAUAAATGCTAAATTAATTTCGCAATCTGTCTGGAAGTTA	76419
286	286
76418	ATTGCTTGCTTTTGTTCATTATTTCTGTACTGTGTGGATCTTCTTAA	76369

286Cys... 286

76368 AGGTAATTTGGATTAGAGAAACCGATGCATAATCATATGCGCTAACCTCTC 76319

287 287

76318 TATTGTGTTCTATTAGTCTGTAAGCATGATGTTTGGCCCACTCGCTCT 76269

298 oCysProValCysAlaAsnGlyMetTyrTyrAsnGlyGlnTyrGlnC 315

76368 CTGCCCAATGCTCTGGGCATCTTCTTCTCCGGAGGACTTTACCGAT 76219

315 ySerGlyAsnValSerGluTyrSerLysCysThrTyrSerAlaThrGlu 331

76218 GCCATGATACATCTCAGAAATGGAGCAATGTTCTCATTCACATTTGGAT 76169

332 ProValArgValLysLysLysTyrGlnIleProHisGlyThrLysAsnAs 348

76168 CCAGACCGCATCAAGGGAAGTGGAAATCCCTGCAGNAACAGAAAAATCA 76119

348 pTyrLeuMetLys... 352

76118 ATTCTCTTCAAGGTAAACCAAGTCTTTTGTAGAGTTGATATTGGTATGG 76069

352 352

76068 GTTAGGTTAGATGCTGGTATATATTACTCATCAAGCTTAATATTTATATTT 76019

353TrpPheLysSerGlnLysValLysLysProGluArgVa 365

76018 GATTACTTGCAGTGAATAAGTCTCAAAAGAGTGTGAAGCCAAACGAT 75969

365 lLeu...ProPrometSerProGluLysSer...GlySerLysAlaThrC 380

75968 TCTGGCTCTGTATTGTCTGGCAGACATCTCAGGGTCAAGGTTCTAAAG 75919

380 lnArgThrSerIleuSerSerLysGlyLeuAspLysLeuArgPheSer 396

75918 ATGCAACTGACTCTCAAGGAGTCAAAGGCTCAAGGCTACAGATCTTAAAGTTTCA 75869

397 ValValGlyGlnSerLysGluAlaAla... 405

75868 ATTGCTGGAATACTAAGGAAGGATATGTTCTTCTGTCAAAATTTATAT 75819

406AsnG 407

75818 TGTAGACTATCTTTGCACCTTCATTGCTTGATGGATCATTTTCCACGACAC 75769

407 luTrpIleGluLysLeuLysLeuAlaGlyAlaAsnPheTyrAlaArgVal 423

75768 CATGGAAGAGAGAAATTTGAGGAAGCTGGTGACAGTTCATGCTAATGTT 75719

424 ValLysAspIleAsp... 428

75718 AAAAAGGTATGGATACTCGAGTGCTATTTTGTGATTTATGTGACACTA 75669

429Cys...LeuI 431

75668 GTTTTAGCTTTTGAATATCTCTCTTCACAAATTTGTAGGTACAAGCTGT 75619

431 leAlaCysGlyGluLeuAsp...AsnGluAsnAla 441

75618 TTGGTTGTTTGTGGACTGACAGATATCAGAGACGCTGGAATGAGAAAGGC 75569

442 GluValArgLys... 445

75568 AAGGTAGAATATTGCCCTTCATTGTGCTGATACTCTTTAACTTTGTTCACAC 75519

446AlaArgArgL 449

75518 TATTTTTCACATTGAACGCTGGAGTTGTAATATGTTGGCAGAGGA 75469

449 euLysilleProileValargGluGlyTyrIleGlyGluCysValLysArg 455
::: ||||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
75468 TGAAGTGGCAATCGTGAGAGCATTAATTGGTTGATCTGTAAAAA 75419

456 ThrLysCysCysHisLeulleCysIleAsnTrpAsnAlaLeuGluSerSe 482
::: ::: ::||| ::||| ::||| ::||| ::||| ::||| ::|||
75418 CAGAGG.....AAACTTCCATTTGACAAGTPACAAAATTGAAGACAC 75378

482 rLysGly***ThrValThrValLysValLysGlyArgSerAlaCysSer* 499
::: ||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
75377 TAGTCAGAGCCCTTGTCACCTGTAAAGTAAAGACGACGAGCGCT..... 75335

499 **SerPro**ValCysLysAsnThrAlaHisIlePro*****TrpGlu 515
::: ::||| ::||| ::||| ::||| ::||| ::||| ::|||
75334GTGCATGAAGCGTCTGGCGCTCCAA 75311

516 LysHisIleGlnCys**Leu.....LysHi 524
::||| ||| ::| ::| ::| ::| ::| ::| ::|
75310 GAGCAC.....TGTCACATTTCTGACATGGGAACAGTAGTATCTATAACAC 75267

524 sValLeuThr***His***ValCysThrGly.....TyrrYrValL 538
::|||:: ::|||:: ::|||:: ::|||:: ::|||:: ::|||:: ::|||::
75266 RACTCTGACATGCTGTATCTCTACCGGTATCAATAGTTATTACATAC 75217

538 euGlnIleleGlucInAspSpGlySerGluCysTyrrValPheArdLys 554
::|||:: ::|||:: ::|||:: ::|||:: ::|||:: ::|||:: ::|||::
75216 TCCAGATAATCCAAGAAGATAAAGGTTCCAGATTGTTACGTTATTCGTAAA 75167

555 TrpGlyArgValGlySerGluLysIleGlyGlyGlnLysLeuGluMe 571
::|||:: ::|||:: ::|||:: ::|||:: ::|||:: ::|||:: ::|||::
75166 TGGGCCCGAGTTGGAATGAAAGATTGGTGGTAACAAAGTGGAGGAAT 75117

571 tSerLysThrGluAlaIleLysGluPheLysArgLeuPheLeuGluLysT 588
::|||:: ::|||:: ::|||:: ::|||:: ::|||:: ::|||:: ::|||::
75116 GTCAAAGTCTGATCGGTTCCAGAAATCAAACGCTATTCTTCGAAAAA 75067

588 hrGlyAsnSerTrpGluAlatrpGluCysLysThrsnPheAryLysGln 604
::|||:: ::|||:: ::|||:: ::|||:: ::|||:: ::|||:: ::|||::
75066 CCGGAAACACATGGGAATCTTGGGACAAAAAACGAATTTCCAGAAACAA 75017

605 ProGlyArgPheTyrrProLeuAspVal..... 613
|||||::||| |||||::|||
75016 CCTGGAAAATTCTCCCGTTGGACATTTGAAGCATCCATATGACTATATC 74967

613 613

74966 TTGTGCTCTTAAGATTCCTTATGATCGTTCAATAAACTCTTTGTATC 74917

614 AspTyrrGlyValLysLysAlaProLysArgLysAspIle 626
|||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
74916 GTTTGTTTATGAGTATTGGAGTTAATAGCAAGTAGCCAAAAAGAGCCA 74867

627 SerGluMetLysSerSerLeuAlaProGlnLeuLeuGluLeuMetLysMe 643
::: ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
74866 TTTTCAGACAGTAGCAACCTGCTCCATCATTAATAGAAATTGATGAAGAT 74817

643 tLeuPheAsnValGluThrTyrrArg..Ala..... 652
|||||::|||::|||::|||::|||::|||::|||::|||::|||::|||
74816 GCTTTTGTATGGAAACTTACAGGTTAGAACTTATTTGGCATTTGCTT 74767

652 652

74766 CTTCCTCTTCTCTCCCTAGAACTGTATGAAGTAACCTAATAATTACT 74717

653 AlaMetMetGluPheGlu***AsnMetSerGluMetPro 665
|||||::|||::|||::|||::|||::|||::|||::|||::|||::|||
74716 TC CGCAGATCTGCAATGATGGAGTTTCGAGATAAATATGTCAGAGATGCCA 74667

666 LeuGlyLysLeuSer*****AsnIleGlu... 676
|||||::|||::|||::|||::|||::|||::|||::|||::|||::|||
74666 CTTGGGAAGCTCAGCAAAACATAATACAGAAAGGGAAGTAACTCTACTTCTC 74617

676 676

[illegible]

73722 ACCAAGGACTGAGAAATTCACCTCCAGAGCTCTCTGCTACTGTTACATG 73673
859Met...PheGlyLysGlyLeuTyrPheAlaAspLeuValSerLysS 874
73672 GTCAGTACTCTATCTGCTCTTATTTGTGGGATCTTATCAATTGAGACT 73623
859Met...PheGlyLysGlyLeuTyrPheAlaAspLeuValSerLysS 874
73622 TCTTCACTATCTAGTTAACTTGTGTAGAGAATCTCAACTTTTTTAAACCC 73573
860Met...PheGlyLysGlyLeuTyrPheAlaAspLeuValSerLysS 874
73572 CATGTTTCAGTTTGGAAAGGATATTAATCTTGTGACCTTGTGACGATAAA 73523
874 erAlaGlnTyrCysTyrValAspArgAsnProValGlyLeuMetLeu 890
73522 GTGCTCACTACTGCTACACTGTGAAGAAAATCCGGTGGGTCTAATGCTT 73473
891 LeuSerGluValAlaLeuGlyAspMetTyrGluLeuLysAlaThr... 906
73472 CTGAGTGAAGTTCATTCGGGAGAAATACATGACTAACAAAAGCTAAGCT 73423
906Met...PheGlyLysGlyLeuTyrPheAlaAspLeuValSerLysS 874
73422 GAGCCCATATTACAGTTTTCAGGTTATGCTACTAACCATTGGTTCCTTA 73373
906Met...PheGlyLysGlyLeuTyrPheAlaAspLeuValSerLysS 874
73372 TCTTGAACACTATATGAACAACTTAAGATATTACTTACCTTCATATATGA 73323
907Met...PheGlyLysGlyLeuTyrPheAlaAspLeuValSerLysS 874
73322 CTTTCTCTATGCTCCTCAATCTTTAGTATATGATAAACCCTCCGAGAG 73273
914 yLysHisSerThrLysGlyLeuGlyLysThrValProLeuGluSerGluP 931
73272 GAACACTCGACCAAGGCTCGACRAAGAGTCCCTCAAGATCCGAGT 73223
931 heValLysTrpArgAspValValProCysGlyLysProValPro 947
73222 TTGCCAAGTGGAGAGTGATGCTGACTTTCCTCTGTGAAAACCTGTTCAT 73173
948 SerSerLysSerGluLeuMetTyrAsnGluTyrIleValTyrAs 964
73172 TCAAAGTCAAGGCTCTGAGCTTATGTACAATGAGTATATGCTTACGA 73123
964 nThrSerGlnVal.....LysMetGlnPheLeuLeu 974
73122 TACAGCCAGGTAACCTGTGCTTTCCAAATAATCTTCCTTGT 73080
seq_name: gb_ro:AF168781
seq_documentation_block: 3047 bp mRNA ROD 28-JUL-1999
LOCUS AF168781 Cricetus griseus poly ADP-ribose polymerase (PARP) mRNA, complete cds.
DEFINITION
ACCESSION AF168781
VERSION AF168781.1 GI:5616519
KEYWORDS
SOURCE Chinese hamster.
ORGANISM Cricetus griseus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Cricetinae; Cricetulus.
REFERENCE 1 (bases 1 to 3047)
AUTHORS Ganesh.A.N., Phillips,E., Shah,R., Affar,E.B., Poirer,G., Thacker,J. and Meuth,M.
TITLE Suppression of the radiation sensitive phenotype of hamster 1rs1 and 1rs2 strains selected for resistance to 3 aminobenzamide unpublished.
JOURNAL 2 (bases 1 to 3047)
REFERENCE 2 (bases 1 to 3047)
AUTHORS Ganesh.A.N. and Meuth,M.
TITLE Direct Submission

Mon Mar 11 09:58:37 2002

6 LysAlaThrLysAlaGluTyrAlaLysSerGlyArgAlaSerCysLys 22
114 AAGCTCTATGAGTCAGTACGCAAGAGCGAGCGCCCTCTTGCAGAA 163
22 rCysArgSerProIleAlaLysAspGlnLeuArgLeuGlyLysMetVal 39
164 ATGACGAGGAGATCCCAAGGACTCGCTCCGATGGCCATCATGGTGC 213
39 InAlaSerGlnPheAspGlyPheMetProMetTrpAsnHisAlaArgCys 55
214 AGTCGCCCATGTTGTATGGAAGAGTCCACACTGGTACCACCTTCTCTGC 263
56 IlePheSerLysLysAsnGlnIleLysSerValAsp... AspValGlu 71
264 TTCTGGAAGTGGGCCCATCCATCCGCCACCCCTGACGTTGAGGTGGATG 313
71 YlleAspAlaLeuArgTrpAspGlnGluLysIleArgAsnTyrVal. 87
314 GTTCTCTGAGCTTCGGTGGATGACACAGCAGAAAGTCAAGAACACACGG 363
88GlySerAlaSerAlaGlyThrSerSerThrAlaAlaProPro 101
364 AAGCTGGAGGAGTACAGCAAGGCGGAGGATGGAATTGGTAGCAAGGCA 413
102 GluLysCys.....ThrIleGluIleAlaProSerAlaArgTh 114
414 CAGAAGACTCTGGGTGACTTTGACAGCAGATGATGCCAAGTCCACACAGA 463
114 rSerCysArgArgCysSerGluLysIleThrLysGlySerValArgLys 131
464 TACGTGCAAGGGGTGTATGGAAGAGATAGAAAGGCGCAGGTGGCCCTGT 513
131 erAlaLysLeu...GluSerGluGlyProLys...GlyIle.....Pro 143
514 CCAGAAGATGTGGACCGGAGAGGACACAGCTAGGATGATGACCGC 563
144 TrpTyrHisAlaAsnCysPhePheGlu.....ValSe 154
564 TGGTACCATCAGGCTGCTTTGTCAACAGCAGGAGGAGTGGTTCGG 613
154 rProSerAlaThrValGluLysPheSerGlyTrpAspThrLysSerAsp 171
614 GCCCGAGTACAGTGGAGTACAGTCAAGGGCTTCAGCTCTCTGCTACAG 663
171 LuAspLysArgThrMetLeuAspLeuValLysLysAsp.....Val 184
664 AGGATAAA.....GAAGCCCTGAAGAAGCAGCTCCAGGAGTC 701
185 GlyAsnAsnGluInAsnLysGlySerLys..... 194
702 AAGAGTGAAGGAAGAGAAAGAGCGGATAGGTGGATGGAGTGAAGT 751
195ArgLysLysSerGluAsnAspIleAspSerTyrLysSerAlaArg 210
752 GCGCAAGAGAAATCTAAAGAGAAAGAGAC.....AAGATAGTAGC 795
210 euAspGluSerThrSerGluGlyThrValArgAsnLysGlyInLeuVal 226
796 TTGAAAAA..... 803
227 AspProArgGlySerAsnThrSerSerAlaAspIleGlnLeuLysLeu 243
804CGCTAAA 811
243 sGluInSerAspThrLeuTrpLysLeuLysAspGlyLeuLysThrHis 260
812 GCTCAGACGACCTGATCTGGAACATCAAGCAGCAGCTTAAAGAGTGT 861
260 alSerAlaAlaGluLeuArgAspMetLeuGluAlaAsnGlyGlnAspThr 276
862 GTTCAACTAATGACCTGAGGAGCTACTCATCTTCAACAGCAGCAAGTG 911
277 SerGlyProGluArgHisLeuLeuAspArgCysAlaAspGlyMetLeuPh 293

912 CTTCTGGGAGTCGGCATCTTGGACCGAGTAGCTGATGCATGGTGT 961
293 eGlyAlaLeuGlyProCysProValCysAlaAsnGlyMetTyrTyr 310
962 CGGTGCCCTCTTCCTCGCAGGAATGCTCGGTGAGCTGTCTTCAAGA 1011
310 snGlyGlnTyrGlnCysSerGlyAsnValSerGluTrpSerLysCysThr 326
1012 GCGATGCTTATCTGCTACCTGGGAGCTGCTGCTGACCAAGTGTATG 1061
327 TyrSerAlaThrGluProValArgValLysLysLysTrpGlnIleProHi 343
1062 GTCAAGACACACACACCAACCG.....AAGGAGTGGTAAACCCA... 1103
343 sGlyThrLysAsnAspTyrLeuMetLysTrpPheLysSerGlnLysVal 360
1104AAGGAATCCGAGAAATCTTACCTCAAGAAATGAAGGTTA 1146
360 yLysProGluArgValLeuProProMetSer..... 370
1147 AAAGCAGACCGTATATTCCTCCAGAACACCGCTCCGTGGCCAC 1196
371ProGluLysSerGlySerLysAlaThrGlnArgThrSerLeu 385
1197 CTTCCGCCCTCCACAGCTCGGCTCTGCTGCTGAACCTCTCTCTTC 1246
385 uSerSerLysGlyLeuAspLysLeuArgPheSerValValGly..... 399
1247 AGCAGATAAGCCATTATCCCAAGATGAAGATCTGCTCTCGGGAAGCTGT 1296
400 ..GlnSerLysGluAlaAlaAsnGluTrpIleGluLysLeu.....Lys 413
1297 CCGGAACAGGATGAAGTGAAGCCATGATTCAGAACTCGGGGGAAG 1346
414 LeuAlaGlyAlaAsnPheTyrAlaArgValValLysAspIleAspCys 430
1347 TTAACGGGAGCGCCCAACAGGCTTCCTG.....TGCAT 1381
430 uIleAlaCysGlyGluLeuAspAsnGluAsnAlaGluValArgLysAla 447
1382 CAGCACCACAAAGAGGAGTGAAGAGATGAATAAGAGTGGAGGAGTAA 1431
447 rGArgLeuLysIleProIleValArgGluGlyTyrIleGlyGluCysVal 463
1432 AGAAGCCACATCCGAGTGTCTCTGAGGACTCTCCAGGACGCTCC 1481
464 LysArgThrLys.....CysCysHisLeuLysCysIle 474
1482 GCCTCCACCAAGAGCTTCAGGAGTGTCTTAGCGCACATCTGTCC... 1529
474 eAsnTrpAsnAlaLeuGluSerSerLysGly***ThrValThrValLysV 491
1530 .CCTTGGGGGCGAGGTGAAGCAGAGCCTGTGAAGTT...GTGCC 1575
491 alLysGlyArgSerAlaCysSer***SerPro***ValCysLysAsnThr 507
1576 CAAGGGAAGTCAAGGCTCGCTCTCCAAA.....AAAGCAAG 1616
508 AlaHisIlePro****Trp.....GluLysHisIleGlnCys 520
1617 GCCCAGTCAAGGAGGAGGTATCAACAAATCTGAAAAGAGAAATGAAT 1666
520 s***LeuLys..... 523
1667 AACTCTTAAAGAGGAGCAGCTGTGGATCTGATTCGACTGGAACT 1716
524HisValLeu.....Thr***His 529
1717 CTGCGCATGCTCTGAGAAAGGAGTGGAGTCTTCAGTGCACCTTGGC 1766
530 ***Val.....CysThrGlyTyrValLeuGlnIleI 541
541

```

1767 CTGGTGGACATCGTTAAAGGAACCAACTCCTACTACAAGTCGACCTTCT 1816
541 egluGlnaspGspGlySerGluCysTyrValPheArgLysTrpGlyArgV 558
      |||:|||||
1817 GGAGGACACACAGNAACAGGTATTGGATATTCAGGTCCTGGGGCCGTG 1866
558 alGlySerGluLysIleGlyGlnLysLeuGluMet...SerLys 573
      |||:|||||
1867 TGGGTACG...GTGATCGGTAGCAACAACCTGGAACAGATGCCGTCCAAG 1913
574 ThrGluAlaIleLysGluPheLysArgLeuPheLeuGluLysThrGlyAs 590
      |||:|||||
1914 GAGATGCCATGAGCACTTCATCAATATATATGAAGAAAAACCGGAA 1963
590 nSerTrpGluAlaTrpGluCysLysThrAsnPheArgLysGlnProGlyA 607
      |||:|||||
1964 CGCTTGGCACTCC.....AAAATTTTCCACCAAGTATCCCAAAA 2001
607 rGpHeTyrProLeuAspValAspTyrGlyVal...LysLysAlaProLys 622
      |||:|||||
2002 AGTTCTACCCCTGGAGATTGACTATGGCCAGGATGAAGAGCAGTGAAG 2051
623 ArgLysAspIleSerGlu...MetLysSerSerLeuAlaProGlnLeuLe 638
      |||:|||||
2052 AGCTGACAGTAAATCTCGCACCAAGTCCAAGCTCCCAAGCCAGTTCA 2101
638 uGluLeuMetLysMetLeuPheAsnValGluThrTyrArgAlaAlaMetM 655
      |||:|||||
2102 GGACCTCATCAAGATGATCTTTGATGGAAAGTATGAAGAAGCCATGG 2151
655 etGluPheGlu***AsnMetSerGluMetProLeuGlyLysLeuSer*** 671
2152 TGGAGTATGAGATCGACCTTTCAGAAGATGCCCTTGGGAAGCTGAGCAA 2201
672 ***AsnIleGlu***GlyPheGluAlaLeuThr*****LeuPh 688
2202 AGCAGATCAGCGCCGATCTCCATCCTCAGT..... 2234
588 eGluGlyHisArg***SerSerThrGlyLeu***GluLysAla***LeuL 705
      |||:|||||
2235 .GAGTCCAGCAGCGCGGTGTCTCAGGCGACGACGACTCTCAGATCCTGG 2283
705 eu*****PheSerLeuLeuSerLeuLeuPheIleLeuIle 721
      |||:|||||
2284 ATCTCTCAATCCTTTTACACCTCATGCCCCACGACTTT..... 2324
722 LeuTyrGlyMetArgMet.....IleSerTy 730
2325 .....GGGATGAAGAAGCTCGCTCCTGAACAATGCAGACAGTGC 2368
730 rSerLysAlaLysMetLeuGluAlaLeuGlnAspIleGluIleAlaSerL 747
      |||:|||||
2369 GGCCAGGTGGAAATGCTTGACAACTGCTGGACATCGAGGTGGCTACA 2418
747 yIleVal.....GlyPheAspSerAspSerAspGluSerLeuAspasp 761
      |||:|||||
2419 GTCTGCTCAGGGAGGCTGTGATGATAGCAGCAAGATCCCATCGATGTC 2468
762 LysTyrMetLysLeuHisCysAspIleThrProLeuAlaHisAspSerGl 778
      |||:|||||
2469 AACTATGAGAAGCTCAAACTGACATTAAGTGGTTGACAGAGATTCGA 2518
778 uAspTyrLysLeuIleGluGlnTyrLeuLeuAsnThrHisAlaProThrH 795
      |||:|||||
2519 AGAGCCGAGATCATCAGGAATATGTTAAGAACACTCATGCAACACAC 2568
795 iLysAspTrpSerLeuGluLeuGluValPheSerLeuAspArgAsp 811
      |||:|||||
2569 ACAATCGTATGACTTGAAGTCATCGATATCTTTAAGATAGAGCGTGA 2618
812 GlyGluLeuAsnLysTyrSerArgTyrLysAsnLeuHisAsnLysMe 828
      |||:|||||
2619 GCGGAATGCCAGCGTTAAGACCCCTTTAAG...CAGCTTCAACACCGAAG 2665
828 tLeuLeuTrpHisGlySerArgLeuThrAsnPheValGlyIleLeuSerG 845
      |||:|||||
2666 ATTGCTGTGGCAGGGTCCAGGACCACTTTGCTGGGATCCTGCTCC 2715
845 lncGlyLeuArgIleAlaProGluAlaProValThrGlyTyrMetPhe 861
      |||:|||||
2716 AGGCTCTTCGGATAGCCCGCTGAAGCCCGCTGACAGCGCTACATGTT 2765
862 GlyLysGlyLeuTyrPheAlaAspLeuValSerLysSerAlaGlnTyrCy 878
      |||:|||||
2766 GGTAAAGGATCATTTCTGCTGACATGGTCTCAAGAGTGCCCACTACTG 2815
878 sTyrValAspArgAsnAsnProValGlyLeuMetLeuLeuSerGluValA 895
      |||:|||||
2816 CCATAGCTCTCAGGGAGACCCCAATAGGCTTAATCTGTGGGAGATG 2865
895 lalGlyAspMetTyrGluLeuLysLysAlaThrSerMetAspLysPro 911
      |||:|||||
2866 CCCTTGGAAACATGTATGAATGAAGCAGCGCTTCACATATCAGCAAGTTA 2915
912 ProArgGlyLysHisSerThrLysGlyLeuGlyLysThrValProLeuGl 928
      |||:|||||
2916 CCCAAGGCAAGCACAGTCTCAAGGTTTGGCAAACTACCCCTGATCC 2965
928 uSerGluPheValLysTyrArgAspValValValProCysGlyLysP 945
      |||:|||||
2966 TTCAGCTAACATTAGT...CTGGATGGGTAGACGCTCTCTTGGGACCG 3012
945 roValProSerSerIleArgSerSerGluLeuMetTyrAsnGluTyrIle 961
      |||:|||||
3013 GGATTTTCATCTGGTGTGAATGACACCTCTCTACTATATACGAGTACATT 3062
962 ValTyrAsnThrSerGlnValLysMetGlnPheLeuLysValArgph 978
      |||:|||||
3063 GCTCATGATATTGCTCAGTAAATCTGAAGTATCTGCTGAAACTGAAATT 3112
978 ehIshisLys 981
      |||:|||||
3113 CAATTTTAAG 3122
seq_name: gb_pr:HUMPOLP
seq_documentation_block:
LOCUS HUMPOLP 3640 bp mRNA PRI 08-JAN-1995
DEFINITION Human poly(ADP-ribose) polymerase mRNA, complete cds.
ACCESSION M18112
VERSION M18112.1 GI:190166
KEYWORDS polymerase.
SOURCE Human SV40 transformed fibroblast, cDNA to mRNA, clone pPAP.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Uchida,K., Morita,T., Sato,T., Ogura,T., Yamashita,R., Noguchi,S.,
Suzuki,H., Nyunoya,H., Miwa,M. and Sugimura,T.
TITLE Nucleotide sequence of a full-length cDNA for human fibroblast
poly(ADP-ribose) polymerase
JOURNAL Biochem. Biophys. Res. Commun. 148 (2), 617-622 (1987)
MEDLINE 88076933
COMMENT Draft entry and computer readable sequence for [1] kindly provided
by K.Uchida, 02-MAR-1988.
FEATURES
source
1..3640
/db_xref="taxon:9606"
/map="1q41-q42"
<L..3640
/note="POLP mRNA"
140..3184
/ gene="ppol"
140..3184
/ gene="PPOL"

```


870 TCTGCTTCCCTTGAGGAGGTGTAAGGGCAGTTTGTGTTCAAGACGCATG 919
 312 lntYrGlnCysSerGlyAsnValSerGluTrpSerLysCysThrTySer 328
 920 CATACTACTCTCAGGGGACATTACTGCTGGACTAAGTGTGGCTGCTAAA 969
 329 AlaThrGluProValArgValLysLysTrpGlnLeProHisGlyTh 345
 970 ACACAGACTCCCAATAGG.....AAGACTGGGTAAATCCCG..... 1005
 345 rLysAsnAspTyLeuMetLysTrpPheLysSerGlnLysValLysLys 362
 1006 .AAGGAATTCTGTGAGATTCTTATCTCAAAAATTTAAGTGTAAAGAGC 1054
 362 roGluArgValLeuProMetSer.....Pro 371
 1055 AGGATAGGATATTCCTCTCTCAGGCTGCACTGTGAACCTGCACACCAT 1104
 372 GluLysSerGlySerLysAlaThrGlnArgThrSerLeuLeuSerLys 388
 1105 CCACCTGCATCCGCTCCATTGACACAGACTGCTCACTGCACCCCAAGCAA 1154
 388 sGlyLeuAspLysLeuArgPheSerValValGly.....GlnSerL 402
 1155 ACCACTGCACCAACATCAAAATCTGACTCTTGGAAAGCTGTCCAAGAACA 1204
 402 ySGLuAlaLaAsnGluTrpIleGluLysLeuLysLeuAlaGlyAlaAsn 418
 1205 AGGAGGAAGTGAAGAATAATTGTGGAGGAGCTG.....GGAGGAAG 1245
 419 PheTyrAlaArgValValLysAspLysAspCysLeuIleAlaCysGly 435
 1246 ATGACAAACAACAGCTAAACAGCCACCCTGTGCATCAGCACACAAAGGA 1295
 435 uLeuAspAsnGluAsnAlaGluValArgLysAlaArgLeuLysLys 452
 1296 GTGGAGAAATGACACAGAAGATGGAGAAGTGAAGATGCCAAGTTC 1345
 452 roLleValArgGluTyTrpIleGlyGluCysValLysArgThrLys... 467
 1346 GTGTGCTCAGAGGAGTTCTTAAAGATGTGAATCTTCCACAAAGGCG 1395
 468CysCysHisLeuIleCysIleAsnTrpAsnAlaLe 479
 1396 TTTCAGGAGCTTCTCTCTCCATCAATTCA...CCTTGGGGTCCACA 1442
 479 uGluSerSerLysGly***ThrValThrValLysValLysGlyArgSerA 496
 1443 GTGTAAACGGAGCACACGAGGAGTGGCGGTGATGGGAAG..... 1482
 496 laCysSer***SerPro***ValCysLysAsnThrAlaHisIle..... 510
 1483 ...TGCAGC...AGCCTCAAAATATGAAGAGTCTGGGAAGTTCACAGAA 1527
 511Pro***TrpGluLysHisIleGlnCys***LeuLys... 523
 1528 GAACAAGGACCTAGCAAGTCTGAGAAGAAATCAAGCTAACAGTTAAGG 1577
 524HisValL 526
 1578 TGGAGCAGCGTAGATCCTGATTCTGGCTTGGAGGATTTCTGCTCATGTCT 1627
 526 eu.....Thr***His***Val..... 531
 1628 TTGAAAAGGTGGAGATTTTCAGTCGCAACTCTGGCCTAGTGATATT 1677
 532CysThrGlyTyTrpValLeuGlnIleIleGluGlnAspAs 545
 1678 GTGAAAGGAACAATTCCTATTACAACATCGAGCTCTGGAGATGACAG 1727
 545 pGlySerGluCysTyValPheArgLysTrpGlyArgValGlySerGluL 562

[illegible]

```

849 leaIaProGluAlaProValThrGlyTyrMetPheGlyLysGlyLeu 865
|||||
2627 TAGTCCCTCCGAGCTCCTGTGACCGGTACATGTTGGGAGGCATC 2676
|||||
866 TyrPheAlaAspLeuValSerLysSerAlaGlnTyrCysTyrValAspAr 882
|||||
2677 TACTTTGAGATATGTTTCCAGAGAGTGCACACTACTGTACACATCTCA 2726
|||||
882 gAsnAsnProValGlyLeuMetLeuSerGluValAlaLeuGlyAspM 899
: |||||
2727 AGCTGATCAATTTGGGTTAATACTACTGGGAGAGTTGCCCTCGGAATA 2776
|||||
899 etTyrGluLeuLysLysAlaThrSerMetAspLysProProArgGlyLys 915
|||||
2777 TGTATGAGCTTAAGATAGCTTCTCACATACAAATATGCCCAAGGAAAA 2826
|||||
916 HisSerThrLysGlyLeuGlyLysThrValProLeuGluSerGluPheVa 932
|||||
2827 CATAGTGTGAAGGCTTGGCAAACTGCACCT...GATCCACACGCCAC 2873
|||||
932 LysTrpArgAspAspValValValProCysGlyLysProValProSerS 949
: |||||
2874 TACAACCTCTGATGCTAGAGTTCCCTTAGGGAATGGGATCTCAACAG 2923
|||||
949 erIleArgSerSerGluLeuMetTyrAsnGluTyrIleValTyrAsnThr 965
|||||
2924 GAATTAATGATGACTGCTGCTGTACAAATGAATATATGTTGATGATGT 2973
|||||
966 SerGlnValLysMetGlnPheLeuLeuLysValArgPheHisLys 981
: |||||
2974 GCTCAGTAAATCTGAAGTACCTGCTGAACTGAAATCAACTATAAG 3021
|||||
seq_name: gb_pr:HUMRISDAD

seq_documentation_block:
LOCUS HUMRISDAD 3795 bp mRNA PRI 09-JAN-1995
DEFINITION Human poly(ADP-ribose) synthetase mRNA, complete cds.
ACCESSION J03473
VERSION J03473.1 GI:337423
KEYWORDS poly(ADP-ribose) synthetase.
SOURCE Human placenta, cDNA to mRNA (library of H.Okayama), clones pPARS[1,11,21,32,32,41,-F].
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 3795)
Kurosaki,T., Ushiro,H., Mitsuuchi,Y., Suzuki,S., Matsuda,M.,
Matsuda,Y., Katunuma,N., Kangawa,K., Matsuo,H., Hirose,T.,
Inayama,S. and Shizuta,Y.
Primary structure of human poly(ADP-ribose) synthetase as deduced
from cDNA sequence
J. Biol. Chem. 262 (33), 15990-15997 (1987)
88058958
Draft entry and printed copy of sequence for [1] kindly provided by
Y.Shizuta, 23-NOV-1987.
FEATURES
source
1..3795
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="1q41-q42"
<1..3795
/notes="poly(ADP-ribose) synthetase mRNA"
96..3140
/gene="PPOL"
96..3140
/notes="poly(ADP-ribose) synthetase"
/codon_start=1
/db_xref="GDB:G00-119-508"
/protein_id="AAB59447.1"
/db_xref="GI:337424"
/translation="MAESSDKLYRVEYAKSERASCKKCSIPKDSLRMAIMVQSPMF

```

old_sequence

```

804..805
/ gene="ppol"
/ note="cg"
/ citation=[1]
/ replace=""
1193..1199
/ gene="ppol"
/ note="ccac"
/ citation=[1]
/ replace=""

BASE COUNT 1048 a 847 c 1037 g 863 t
ORIGIN 73 bp upstream of AluI site.

```

alignment_scores:

```

Quality: 1593.00 Length: 1071
Ratio: 2.309 Gaps: 37
Percent Similarity: 64.426 Percent Identity: 36.321

```

alignment_block:

```

US-09-236-995D-2 x HUMRISDAD ..
Align seg 1/1 to: HUMRISDAD from: 1 to: 3795

6 LysAlaTriPysAlaGluTyrAlaLysSerGlyArgAlaSerCysLys 22
|||||
114 AAGCTCTATCGATCGAGTACGCCAAGAGCGAGCGGCTCTTGGCAAGAA 163
|||||
22 rCysArgSerProIleAlaLysAspGlnLeuArgLeuGlyLysMetVal 39
|||||
164 ATGCAGCGAGAGCATCCCAAGGACTCGCTCCGATGGCATCATGGTGC 213
|||||
39 InAlaSerGlnPheAspGlyPheMetProMetTrpAsnHisAlaArgCys 55
|||||
214 AGTCGCCCATGTTTGTGTAAGAAAGTCCACACTGGTACCACCTTCTCTGC 263
|||||
56 IlePheSerLysLysAsnGlnIleLysSerValAsp...AspValGlu 71
|||||
264 TTCTGGAAGTGGCCACCTCCATCCGCGACCTGACGTTGAGGTGGATG 313
|||||
71 yIleAspAlaLeuArgTrpAspAspGlnGluLysIleArgAsnTyrVal.. 87
|||||
314 GTTCTCTGAGCTTCGGTGGGATGACCCAGCAAGTCAAGAACAGACAGCGG 363
|||||
88 .....GlySerAlaSerAlaGlyThrSerSerThrAlaAlaProPro 101
|||||
364 AAGCTGGAGAGTGCACAGGCAAGGCCAGGATGGATTTGATGACAGGCA 413
|||||
102 GluLysCys.....ThrIleGluIleAlaProSerAlaArgTh 114
|||||
414 GAGAGACTCTGGTGACTTTGGCAGAGATATGCCAAGTCCCAAGAGAG 463
|||||
114 rSerCysArgArgCysSerGluLysIleThrLysGlySerValArgLeu 131
|||||
464 TACGTGCAAGGGTGTATGGAGAAGATAGAAAAGGGCCAGGTGCGCTGT 513
|||||

```

```

DGKVPWHYHFCFWKVGHSIRHPDVEVDFSELRWDDQOKVKKTAEGAGVTCKGODGI
GSAKRETLGDFAAEYAKSNRSTCKGMEKIEKQVRLSKNMVDPEKPOLGMDRWHYP
GCFVKNEELGERPEYSASQKLGFSLLATEDKEALKQLPGVKRGRKQKVDGVE
VAKKSKKEDKDKLEKALQNDLIWNIKDELKVCSTNDLKELLIFNKQVPSGE
SATLDRVADGMVFGALLPCEEGSQLVFKSDAYICTGDTATWTKCMKVQTQPNRKEW
TPKEPREISLVKLVKKVKKODRIPPPETSASVAATPPSTASAPAAVNSSASADKPLSN
MKILITGKLSRNKDEKVMIEKLGKLTGTANKASLCISTKKEVKMNKMEVEKEAN
IRVSEDFLDQVDSASTSKQELFLAHILSPWGAEVKAEPEVVAAPRGKSGAALSXSK
GOVKEGINKSEKMKLITLGGAAVDPDLSGSAHVLEKGGKVFATLGLDIVKGT
NSYKQLLEDDEKRNRYWIFRSGRVGTWIGSNKLEQMPKEDAEHFMKLYEEKTGN
AWHSNKFTYKPKFYPLEIDYGODEBAVKLTVPNGTKSLPKPKVODLLKMFIVESM
KKAMEYEIDLOKMPGLKSKROIQAAYSTLSEVOQAVSGGSDSOILDLSNRFYTLI
PHDGMKPPLLNNAVSQAKVEMLONLDIEVAYSLRGGSDSSKDPIDVNYEKLK
TDIKVDRDSEEAELIRKRYVKNTHAIDLEVIDIEFKIEREGECORIKPKQLHN
RRLIWHGSRITNEFAGLSGLRIAPPEAVPTGYMFGKGYIFADVMKSNANYCHTSDG
PIGLILGEVALGNMYELHSHISKLPKHSVKGKLTGTPDPSANISLDGVDVPLG
TGISGVNDTSLLYNEVYDIAQVNLKYLKLFNFKTSLW"

```

```

131 erAlaLysLeu...GluSerGluGlyProLys...GlyIle.....Pro 143
132  |||||:  ::  |||  |||||:  |||||:
514 CAAAGAGATGGTGGACCCGAGAGAGCCACAGCTAGGTCATGATTGACCGC 563
144 TrpTyrHisAlaAsnCysPhePheGlu.....Valse 154
145  |||||:  ::  |||||:  ::
564 TGGTACCATCCAGCTGCTTTGTCAAGAACACAGGAGGAGCTGGGTTCGG 613
154 rProSerAlaThrValGluLysPheSerGlyTyrTrpAspThrLeuSerAspG 171
155  |||||:  ::  |||||:  ::  |||||:  |||||:  |||||:  |||||:
614 GCCGAGTACAGTCGAGTCAAGGCTTCAGCTCCTTGGCTACAG 663
171 luAspLysArgThrMetLeuAspLeuValLysLysAsp.....Val 184
172  |||||:  ::  |||||:  ::  |||||:  |||||:
664 AGGATAAA.....GAAGCCCTGAAGAGCAGCTCCAGCAGATC 701
185 GlyAsnAsnGluGlnAsnLysGlySerLys..... 194
195  |||||:  ::  |||||:  |||||:  |||||:
702 AAGAGTCAAGGAAGAGAGAAAGCGATAGGTGGATGGATGAAGT 751
195 ....ArgLysLysSerGluAsnAspIleAspSerTyrLysSerAlaArgL 210
196  |||||:  ::  |||||:  |||||:  |||||:  |||||:
752 GCGCAAGAGAAATCTAAAGAAAGAAAGAC.....AAGGATAGTAAGC 795
210 euAspGluSerThrSerGluGlyThrValArgAsnLysGlyGlnLeuVal 226
211  |||||:  ::  |||||:  |||||:  |||||:
796 TTGAAGAAAGCC..... 806
227 AspProArgGlySerAsnThrSerSerAlaAspIleGlnLeuLysLeuL 243
228  |||||:  ::  |||||:  |||||:  |||||:
807 .....CTAAA 811
243 sGluGlnSerAspThrLeuTrpLysLeuLysAspGlyLeuLysThrHisV 260
244  |||||:  ::  |||||:  |||||:  |||||:
812 GGTCTAGACGACCTGATCTGGAACTCAAGGACGAGCTAAAGAAAGTGT 861
260 alSerAlaLagLeuArgAspMetLeuGluAlaAsnGlyGlnAspThr 276
261  |||||:  ::  |||||:  |||||:  |||||:
862 GTTCAACTAATGACCTGAAGGAGCTACTCATCTCAACAAGCAGCAAGTG 911
277 SerGlyProGluArgHisLeuAspArgCysAlaAspGlyMetLeuPh 293
278  |||||:  ::  |||||:  |||||:  |||||:  |||||:
912 CCTCTGGGGAGTCGCGCATCTTGGACCGAGTAGCTGATGGATGGTGT 961
293 eGlyAlaLeuGlyProCysProValCysAlaAsnGlyMetTyrTyrTyra 310
294  |||||:  ::  |||||:  |||||:  |||||:  |||||:
962 CGGTGCTCTCTCCCTGGAGGATGCTCGGGTCAGCTGGTCTTCAAGA 1011
310 snGlyGlnTyrGlnCysSerGlyAsnValSerGluTrpSerLysCysThr 326
311  |||||:  ::  |||||:  |||||:  |||||:  |||||:
1012 GCGATGCTTATTACTGCATGGGGAGCTCCTGCGGACCAAGTGTATG 1061
327 TyrSerAlaThrGluProValArgValLysLysLysTyrTrpGlnIleProH 343
328  |||||:  ::  |||||:  |||||:  |||||:  |||||:
1062 GTCAAGACACACACACCCCAACCGG.....AAGGAGTGGTAAACCCCA... 1103
343 sGlyThrLysAsnAspTyrLeuMetLysTrpPheLysSerGlnLysValL 360
344  |||||:  ::  |||||:  |||||:  |||||:  |||||:
1104 .....AAGGAATCCGAGAATCTCTTACCTCAAGAAATGAAGTTA 1146
360 ysLysProGluArgValLeuProMetSer..... 370
361  |||||:  ::  |||||:  |||||:  |||||:  |||||:
1147 AAAAGCAGGACCGTATATTCCCCCAAGAACCCAGCGCTCCGTCGCGGCC 1196
371 .....ProGluLysSerGlySerLysAlaThrGlnArgThrSerLe 384
372  |||||:  ::  |||||:  |||||:  |||||:  |||||:
1197 AGCCCTCCGCTCCACAGCTCGGCTCCTGCTGTGTAACCTCCCTCTGC 1246
384 uLeuSerSerLysGlyLeuAspLysLeuArgPheSerValValGly.... 399
385  |||||:  ::  |||||:  |||||:  |||||:  |||||:
1247 TTCAGCAGATAGCATATATCAACATGAAGATCTGACTCTCGGGAAGC 1296
400 .....GlnSerLysGluAlaAlaAsnGluTrpIleGluLysLeu.... 412

```

```

1297 TGTCGGGAACAAGGATGAAGTGAAGGCCATGATTGAGAAACTCGGGGG 1346
1298  |||||:  ::  |||||:  |||||:  |||||:
413 LysLeuAlaGlyAlaAsnPheTyrAlaArgValValLysAspIleAspCy 429
1340  |||||:  ::  |||||:  |||||:  |||||:
1347 AAGTTAACGGGACGCCCAACAAGGCTTCCCTG.....TG 1381
429 sleuIleAlaCysGlyGluLeuAspAsnGluAlaGluValArgLysA 446
1382  |||||:  ::  |||||:  |||||:  |||||:
1382 CATCAGCACCAAAAAAGAGGTGGAAGAAGTGAATAAGAAGATGGAGGA 1431
446 laArgArgLeuLysIleProIleValArgGluGlyTyrIleGlyGluCys 462
1432  |||||:  ::  |||||:  |||||:  |||||:
1432 TAAAGGAAGCAACATCCAGTGTGTGTGAGGACTTCTCCAGGACGTC 1481
463 ValLysArgThrLys.....CysCysHisLeuIleCy 473
1482  |||||:  ::  |||||:  |||||:  |||||:
1482 TCCGCTCCACCAAGAGCCTTCAGGAGTGTCTTAGCGCATCTTGTGTC 1531
473 sIleAsnTrpAsnAlaLeuGluSerSerLysGly***ThrValThrValL 490
1532  |||||:  ::  |||||:  |||||:  |||||:
1532 C...CCTTGGGGGCGAGAGGTGAAGGCAGAGCCTGTTGAAGTT...GTGG 1575
490 ysValLysGlyArgSerAlaCysSer***SerPro***ValCysLysasn 506
1576  |||||:  ::  |||||:  |||||:  |||||:
1576 CCCCAGAGGGAAGTCAGGGGCTGCGCTCCAAA.....AAAAGC 1616
507 ThrAlaHisIlePro****Trp.....GluLysHisIleGl 519
1617  |||||:  ::  |||||:  |||||:  |||||:
1617 AAGGCCAGGTCAAGGAGGAAGGTATCAACAAATCTGAAAAGAGCAATCAA 1666
519 nCys***LeuLys..... 523
1667  |||||:  ::  |||||:  |||||:  |||||:
1667 ATTAACCTCTAAAGGAGGAGCAGCTGTGGATCTCTGGACTGGAAC 1716
524 .....HisValLeu.....Thr*** 528
1717  |||||:  ::  |||||:  |||||:  |||||:
1717 ACTTGGCCATGTCTGGAGAAAGGTGGGAAGGTCTTCAGTGGCCACCTT 1766
529 His***Val.....CysThrGlyTyrTyrValLeuGlnI 540
1767  |||||:  ::  |||||:  |||||:  |||||:
1767 GGCCTGTGGAGCATCTGTTAAAGGAACCAACTCTACTACAAGCTGCAGT 1816
540 eIleGluAlaAspAspGlySerGluCysTyrValPheArgLysTrpGlyA 557
1817  |||||:  ::  |||||:  |||||:  |||||:
1817 TCTGGAGGACACAAGAAACAGGATATTGGATATTTCAGGTCCTGGGGCC 1866
557 rgValGlySerGluLysIleGlyGlyGlnLysLeuGluMet...Ser 572
1867  |||||:  ::  |||||:  |||||:  |||||:
1867 GTGTGGGTAGC...GTGATCGGTAGCAACAACATGGGAACAGATGCCGTCC 1913
573 LysThrGluAlaIleLysGluPheLysArgLeuPheLeuGlyLysThrG 589
1914  |||||:  ::  |||||:  |||||:  |||||:
1914 AAGGAGTGCATTCAGTTCATCAATATATATGAAGAAAAAACCGG 1963
589 yAsnSerTrpGluAlaTrpGluCysLysThrAsnPheArgLysLysProG 606
1964  |||||:  ::  |||||:  |||||:  |||||:
1964 GAAGCTTGGCCTCC.....AAAAATTCACGAAGTATCCCA 2001
606 lyArgPheTrpProLeuAspValAspTyrGlyVal...LysLysAlaPro 621
2002  |||||:  ::  |||||:  |||||:  |||||:
2002 AAAAGTTCTACCCCTGGAGATTCATCTATGCCAGGATGAAGAGCAGTC 2051
622 LysArgLysAspIleSerGlu...MetLysSerSerLeuAlaProGlnLe 637
2052  |||||:  ::  |||||:  |||||:  |||||:
2052 AAGAAGCTGACAGTAAATCTCTGCACCAAGTCCCAAGCTCCCAAGCAGT 2101
637 uLeuGluLeuMetLysMetLeuPheAsnValGluThrTyrArgAlaAlaM 654
2102  |||||:  ::  |||||:  |||||:  |||||:
2102 TCAGGACCTCATCAAGATGATCTTTGATGGAAAGTATGAAGAAAGCCA 2151
654 etMetGluPheGlu***AsnMetSerGluMetProLeuGlyLysLeuSer 670
1817  |||||:  ::  |||||:  |||||:  |||||:

```

```
2152 TGGTGGAGTATGAGATGACGACCTTCAGAGATGCCCTTGGGGAAGCTGAGC 2201
671 *****AsnIleGlu***GlyPheGluAlaLeuThr*****Le 687
:::|||||::: |||::: |||:::
2202 AAAGGCGAGATCCAGGCGGATACCTCCATCTCTAGT..... 2237
687 uPheGluGlyHisArg***SerSerThrGlyLeu***GluLysAla**L 704
||| ::::: ||| ||| :::::
2238 ...GAGGTCAGCAGCGGTGTCTCAGGCGAGCAGCGACTCTCAGATCC 2283
704 euLeu*****PheSerLeuLeuSerLeuLeuPheIleLeu 720
|| ::::: ||| ::::: |||
2284 TGGATCTCTCAATCGCTTTACACCTGATGCCACGACTTT..... 2327
721 IleLeuTyrGlyMetArgMet.....IleSe 729
|||||:::
2328 .....GGATGAGAGAGCCTCCCTCTGAACAATGCAGACAGTGT 2368
729 rTyrSerLysAlaLysMetLeuGluAlaLeuGlnAspIleGluIleAlas 746
:::|||||::: ||| ||| ||| ||| |||
2369 GCAGGCCAAGGTGAATGCTTGACAACCTGCTGGACATCGAGGTGGCT 2418
746 erLysIleVal.....GlyPheAspSerAspGluSerLeuAsp 760
:::|||||::: ||| ||| ||| ||| |||
2419 ACAGTCTGCTCAGGGAGGGTCTGATAGCAAGCAAGGATCCCATCGAT 2468
761 AspLysTyrMetLysLeuHisCysAspIleThrProLeuAlaHisAspSe 777
:::|||||::: ||| ||| ||| ||| |||
2469 GTCACATATGAGAGCTCAAAAGTGCATTAAGTGGTTCACAGAGATTC 2518
777 rGluAspTyrLysLeuIleGluInTyrLeuLeuAsnThrHisAlaProT 794
|||||::: ||| ||| ||| ||| ||| |||
2519 TGAAGAAGCCGAGATCATCAGGAAGTATGTTAAGAACACATCATCAACCA 2568
794 hrHisLysAspTyrSerLeuGluLeuGluAlaPheSerLeuAspArg 810
|||||::: ||| ||| ||| ||| ||| |||
2569 CACCAATGCGTATGACTTGAAGTCTCGATATCTTTAAGATAGACCGT 2618
811 AspGlyGluLeuAsnLysTyrSerArgTyrLysAsnLeuHisAsnLy 827
:::|||||::: ||| ||| ||| ||| ||| |||
2619 GAAGCGAATGCCAGGTTACAGCCCTTTAAG...CAGCTTCATACCG 2665
827 smetLeuLeuTrpHisGlySerArgLeuThrAsnPheValGlyIleLeu 844
:::|||||::: ||| ||| ||| ||| ||| |||
2666 AAGATTGCTGTGGCAGCGGTCCAGGACCACTTTGCTGGGATCCTGT 2715
844 erGlnGlyLeuArgIleAlaProProGluAlaProValThrGlyTyrMet 860
|||||::: ||| ||| ||| ||| ||| |||
2716 CCCAGGCTCTTCGATAGCCCGCTGAAGCGCGCTGACAGGCTACATG 2765
861 PheGlyLysGlyLeuTyrPheAlaAspLeuValSerLysSerAlaGlnTy 877
|||||::: ||| ||| ||| ||| ||| |||
2766 TTTGGTAAGGATCTATTTCGCTGACATGCTCCAAAGAGTGCACACTA 2815
877 rCysTyrValAspArgAsnAsnProValGlyLeuMetLeuLeuSerGlu 894
|||||::: ||| ||| ||| ||| ||| |||
2816 CTGCATACGCTCTCAGGAGACCAATAGGCTTAATCTGTGGGAGAAG 2865
894 alaIleLeuGlyAspMetTyrGluLeuLysLysAlaThrSerMetAspLys 910
|||||::: ||| ||| ||| ||| ||| |||
2866 TTGCCCTTGGAAACATGATCACTGAAGCAGCGCTTCACATATCAGCAAG 2915
911 ProProArgGlyLysHisSerThrLysGlyLeuGlyLysThrValProLe 927
|||||::: ||| ||| ||| ||| ||| |||
2916 TTACCAGGCGCAGCAGATGCTCAAGGTTTGGGCAAACTACCCCTGA 2965
927 uGluSerGluPheValTyrArgAspAspValValProCysGlyL 944
||| ::::: ||| ||| ||| ||| |||
2966 TCCCTTCAGCTAACATTACT...CTGGATGGGTAGACGCTTCTCTTGGGA 3012
944 ysProValProSerSerIleArgSerSerGluLeuMetTyrAsnGluTyr 960
::: ||| ||| ||| ||| ||| ||| |||
3013 CCGGCTTTTCATCTGCTGTAACACCTCTCTACTATATAAGAGGATC 3062
```

```
961 lleValTyrAsnThrSerGlnVallysMetGlnPheLeuLeuLysValar 977
|||||::: ||| ||| ||| ||| ||| ||| |||
3063 ATGTCTATGATATTGCTCAGTAAATCTGAAGTATCTGCTGAAACTGAA 3112
```

```
977 gPheHisHisLys 981
|||||::: |||
3113 ATTCAATTTTAAG 3125
```

seq_name: gb_ro:BC012041

seq_documentation_block: 3845 bp mRNA ROD 02-AUG-2001
LOCUS BC012041
DEFINITION Mus musculus, Similar to ADP-ribosyltransferase (NAD+; poly (ADP-ribose) polymerase), clone MGC:6498 IMAGE:2648390, mRNA, complete cds.

ACCESSION BC012041
VERSION BC012041.1 GI:15080597

KEYWORDS MGC.
SOURCE house mouse.

ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 3845)

Direct Submission

TITLE

JOURNAL

Submitted (30-JUL-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgabs-r@mail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

cDNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center

Center code: BCM-HGSC

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>

Contact: villalob@bcm.tmc.edu

Villalob, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia, A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W., Muzny, D.M., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAC Plate: 5 Row: P Column: 3

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 49893.

Location/Qualifiers

1..3845

/organism="Mus musculus"

/db_xref="taxon:10090"

/clone="MGC:6498 IMAGE:2648390"

/tissue_type="Mammary tumor. Metallothionien-TGF alpha model. 10 month old virgin mouse. Taken by biopsy."

/clone.lib="NCI_GCAP_Mam1"

/lab_host="DH10B"

/note="vector: pCMV-SPORT6"

60..3104

/codon_start=1

/product="Similar to ADP-ribosyltransferase (NAD+; poly (ADP-ribose) polymerase)"

/protein_id="AAH12041.1"

/db_xref="GI:15080598"

/translation="MAEASRLRYRVEYAKSGRASCCKSSIPKDSLRMAIMVQSPMF DGKVPWHYHFCFWKVGHSIRQPDVEVDGSELRWDDQKVKYKTAEGAGVAGKQDGS GGAEKTLGDELAEYAKSNRSMCKGLEIKGQMLSKKWDPEKPOLGMDRWYHP TCFVKRDELGFREYASASQLKGFLLSAEDKALKQLPAIKNEGRKRGVDGTD E VAKKSKGKDKDSSKLEKALKAGNELIWNILKELKKACSTNDLKEILLFNQOQVPSG ESAILDRVADGMAGFALLPCKECSGLYFSSADYYICGDTVATWTKCMVKTONSRKEW VTPKEFREISYLKLLKVKKODRIFFPESSAPAPLALPLSVTSAPTAVNSAPADKPLS

CDS

NMKILTLGKLSQNKDKAKAVIEKLGKGLTGSANKASLCISTTKKEVEKMSKKMEVAA
 NVRVCEDFQDYSASTKSQELLSASSWGAIEVPAEPKSAAPSKSK
 GAYKEGVNRSERMKLITLKGAAVDPDSGLESHAVHLEKGGVFATGLVDIVKGT
 NSYVKLQLLEDDESRVWIFRSWGRVGTIGSNKLBOMPSKEDAVEHMKLYEEKTN
 AWHSKNFTYKPKFYPLEIDYGODEEAVKLTVPKCTKSLPKPQOELVGMIFDVESM
 KKALEYEIDLKMPGLKLSRRIOIAAYSTLSEVOQAVSQGSESEILDLNRFYTLI
 PHDFGMKKPPLANNADVQAKVEMLNLDIEVAYSLRRGGSDSDSKDPIDVNYEKLK
 TDIKVDRDSEEAERVKYKNTHTATTHAYDLEVIDIFKIEREGESQRYKPFROLHN
 RLLMHWGSRTNFAGILSGLRIAPPEAVPTGYMFGKGIYFADVMVSKSANYCHTSOGD
 PIGLILGELVALGNMVELHSHASIKLPKSHVKGKLTTPDPSPASITLEGVEVPLG
 TGIPTGVNDPCLLYNEYIYVDIAQVNLKLLKLPKFNKISLW"
 BASE COUNT 1042 a 930 c 1080 g 793 t
 ORIGIN

alignment_scores:
 Quality: 1593.00 Length: 1063
 Ratio: 2.315 Gaps: 37
 Percent Similarity: 64.722 Percent Identity: 36.971

alignment_block:

US-09-236-995D-2 x BC012041

Align seg 1/1 to: BC012041 from: 1 to: 3845

3 AlaProProLySAlaThrLySAlaGluTyRAlaLySserGlyArGAlase 19
 |||
 69 GCCTCGGAGAGGCTTATCGAGTGAGTACGGGAAGACGGCGCGCTC 118
 19 rCysLySserCysArgSerProIleAlaLySAspGlnLeuArgLeuGlyL 36
 |||
 119 TTGCAAGAAATCGACGAGAGTATTCCCAAGGACTCCCTCGCATGCCA 168
 36 ysmetValGlnAlaSerGlnPheAspGlyPheMetProMetTrpAsnHis 52
 |||
 169 TCATGGTGCAGTCACCCATGTTGAGTGGGAAGTCCACACTGGTATCCAC 218
 53 AlaArGcYsIlePheSerLySAsnGlnIleLySserValAsp...As 68
 |||
 219 TTCTCTGCTTCTGGAAGTGGGCGACCTCCATCGGCAGCCTGATGTGA 268
 68 pValGluGlyIleAspAlaLeuArGTrpAspAspGlnGluLySleArgA 85
 |||
 269 GGTGATGGCTTCTCTGAGTGGCTGGGATGATCAGCAGAGGTCAAGA 318
 85 snTyRVal.....GlySerAlaSerAlaGlyThrSerSerThrAla 98
 |||
 319 AGACGGCGGAGGCTGGAGCGTGGCAGCAAGGCCAGGATGGAAGTGGC 368
 99 AlaProProGluLyScys.....ThrIleGluIleAlaProse 111
 |||
 369 GCGAAGCGGAGAGACATTTGGTGACATTTTATAGCGAGTACGCCAAGTC 418
 111 rAlaArgThrSerCysArgCysSerGluLySleThrLySglySerV 128
 |||
 419 CAACAGAGGACATGTCAAGAGGCTGCTGGAGAAGATAGAGAAGGCCCA 468
 128 alArgLeuSerAlaLySLeu...GluSerGluGlyProLyS...GlyIle 142
 |||
 469 TCGCGCTGTCCCAAGAAGATGGTGGATCCAGAGAAGCCACAGCTGGGTATG 518
 143ProTrpTyRHisAlaAsnCysPhePheGlu..... 152
 |||
 519 ATTGACCGCTGGTACCATCCAACITGCTTTGTCAAGAAGCGGAGGAGCT 568
 153ValSerProSerAlaThrValGluLySPheserGlyTrpAspThrL 168
 |||
 569 GGGCTTCGGCGCTGAGTACAGTGGCAGTCACTCAAGGGGCTTTAGCCTCC 618
 168 euSerAspGluAspLySArgThrMet.....LeuAspLeuValLyS 181
 |||
 619 TCTCTGCAAGACACAAAGAAGCTCTGGAAGAAGACGAGCTCCCGGCCATCAAG 668

182 LysAspValGlyAsnAsnGluGlnAsnLySglySerLySargLySlysSe 198
 |||
 669AATGAAGA...AAGAGAAAAGGTGA 691
 198 rGluAsnAspIleAspSerTyRlySserAlaArgLeuAspGluSerThrS 215
 |||
 692 CGAGGTGGAT.....GGAACAG 708
 215 erGluGlyThrValArgAsnLySglyGlnLeuValAspProArgGlySer 231
 |||
 709 ATGAA...GTGGCCAAAAGAAATCTAAG.....AAGGGAAG 743
 232 AsnThrSerSerAlaSpIleGlnLeuLySleuGlnSerAspTh 248
 |||
 744 GACAAGGATAGTAGTAGCTGGAGAGGCCCTCAAGGCTCAGAATGAGCT 793
 248 rLeuTrpLySleuLySAspGlyLeuLySThrHisValSerAlaAlaGluL 265
 |||
 794 GATCTGGAATATCAAGACGAGCTGAAGAAAGCGTGTTCACCAAGGACC 843
 265 euArgAspMetLeuGluAlaAsnGlyGlnAspThrSerGlyProGluArg 281
 |||
 844 TGAAGGAGCTGCTCACTTCAACACGAGCAGCGTGCCTCAGGAGAGTCA 893
 282 HisLeuLeuAspArgCysAlaAspGlyMetLeuPheGlyAlaLeuGlyPr 298
 |||
 894 GCGACTTGGACAGAGTGTCTGACGCGATGGCGTTTGGGCGCTTCTGCC 943
 298 oCysProValCysAlaAsnGlyMetTyRtyrAsnGlyGlnTyRGlnc 315
 |||
 944 CTCAAGGAGTGTTCAGCGCAGCTGTCTTTAAGAGCGACGCTTATTACT 993
 315 ySerGlyAsnValSerGluTrpSerLyScysThrTyRserAlaThrGlu 331
 |||
 994 GTACTGGGATGTCACCTGCTGGACCAAGTGCATGTCAGACACAGAAT 1043
 332 ProValArgValLySlySlyTrpGlnIleProHisGlyThrLySAsnAs 348
 |||
 1044 CCTAGCGGA.....AAGGAATGGGTAACCTCA.....AAGGAAT 1078
 348 pTyRLeuMetLySTrpPheLySserGlnLySValLySlySProGluArgV 365
 |||
 1079 CCGAGAAATATCTACTCTCAAGAAGTTAAAGGTCAAAAACAGACCGAA 1128
 365 alLeuProProMetSerProGluLySserGly.....Ser 376
 |||
 1129 TATTCCTCCAGAAAGCAGCGCCCGCCAGCACCTGCCTGCTCCAGACAAGCC 1178
 377 LysAlaThrGlnArgThrSerLeuLeuSerSerLySglyLeuAspLyS... 392
 |||
 1179 GTCACTCAGCACCCACAGCTGTGAACCTCTCTGCTCCAGACACAAGCC 1228
 393LeuArgPheSerValValGlyGlnSerLySg 403
 |||
 1229 CCTGTCTAAGATGAAGATCCTGACTCTGGGAAGCTCTCCAGACAAGG 1278
 403 luAlaAlaAsnGluTrpIleGluLySleu.....LysLeuAlaGlyAla 417
 |||
 1279 ACCAAGCAAAAGCTGTGATTGAGAACTCGAGGCAAGTTGACAGATCT 1328
 418 AsnPheTyRAlaArgValValLySAspIleAspCysLeuIleAlaCysGl 434
 |||
 1329 GCCAACAAGGCTCCTTG.....TGTATCAGCACTAAAAA 1363
 434 yGluLeuAspAsnGluAsnAlaGluValArgLySAlaArgArgLeuLyS 451
 |||
 1364 GGAGGTGGAGAGATGATGAAGAAGATGGAAGAGTGAAGCGGCCAACG 1413
 451 leProIleValArgGluGlyTyRileGlyGluCysValLySargThrLyS 467
 |||
 1414 TTCAGTGTGTGAGGACTTCTCCAGCAGCTGTCTCCCTCCACTAAA 1463
 468CysCysHisLeuIleCysIleAsnTrpAsnAl 478

2307	CCACTCTCTGAACAACGACGACACGCTGCAGGCCAAGGTGGAGTGCCTAGA	2350
737	uAlaLeuGlnAspIleGluLeuLeaSerLysIleVal.....GlyPheA	752
:		:
2357	CAACCTCTCTGCACATCAGGCTGGCCATATGATCTCTTCACGGGTGGCTCTG	2406
752	spSerAspSerAspGluSerLeuAspAspLysTyrMetLysLeuHisCys	768
: : : : : : : : : :		
2407	ACGACAGCAGCAAGGATCCCATCGCTCACTACGAGAACAACCTCAAACT	2456
: : : : : : : : :		
2457	GACATTAAAGGTGGTTGACAGAGATCTTGAAGAGCGCAGTCATCAGGAA	2506
785	nTyrLeuLeuAsnThrHisAlaProThrHisLysAspTyrSerLeuGluL	802
2507	GTACCTGGAAGACACTCATGCTACACGCACACAGCCTATGACCTCGAAG	2556
802	eUGluGluValPheSerLeuAspArgAspGlyGluLeuAsnLysTyrSer	818
: : : : : : : : : :		
2557	TCATCGATATCTTCAAGATAGAGCGCGAGGGGAGAGCCAGCGCTACAAG	2606
819	ArgTyrLysAsnAsnLeuHisAsnLysMetLeuLeuThrHisGlySerAr	835
: : : : : : : : : :		
2607	CCCTTTCAGG...CAGCTTTCACAACCGGAGCGTCTGTGGCACCGCTCCAG	2653
835	gLeuThrAsnPheValGlyIleLeuSerGlnGlyLeuArgIleAlaProp	852
2654	GACCACCAACTTTGCTGGCATCTCTGCGAGGTCTCGCGATAGCCCCAC	2703
852	roGluAlaProValThrGlyTyrMetPheGlyLysGlyLeuTyrPheAla	868
2704	CTGAAGCGCCTGTGACAGGCTACATGTTGGGAAGGGAATCTACTTTGCC	2753
869	AspLeuValSerLysSerAlaGlnTyrCysTyrValAspArgAsnAsnPr	885
: : : : : : : : : :		
2754	GACATGGTGTCCAAAGTGCAAACTACTGCCCACACATCTCAGGGAGACCC	2803
885	oValGlyLeuMetLeuLeuSerGluValAlaLeuGlyAspMetTyrGluL	902
: : : : : : : : : :		
2804	GATTGCTTAATACTGCTGGGAGAGGTGCCCTTGGAAACATGTATCAAC	2853
902	eUlysLysAlaThrSerMetAspLysProArgGlyLysHisSerThr	918
: : : : : : : : : :		
2854	TCAAGCATCTTTCACATATACGAAGTTACCAAGGCAAGCACACAGTGTC	2903
919	LysGlyLeuGlyLysThrValProLeuGluSerGluPheValLysTyrPar	935
: : : : : : : : : :		
2904	AAAGGTTTGGGAACAACACCCCTGACCCTTCGGCCACCATCAC...CT	2950
935	gAspAspValValProCysGlyLysProValProSerSerIleArgS	952
: : : : : : : : : :		
2951	GGAGGGTGTAGAGGTTCCTACTGGGAACAGGATGCCATCTGGTGTCAACG	3000
952	erSerGluLeuMetTyrAsnGluTyrIleValTyrAsnThrSerGlnVal	968
: : : : : : : : : :		
3001	ACACCTGCCTGTGATAATAGTACATGTCTACGACATTCGTACAGTGTC	3050
969	LysMetGlnPheLeuLeuLysValArgPheHisLysLys	981
: : : : : : : : : :		
3051	AAGTCACAAATCTCTCAAACTCAAGTCAATTTTAAAG	3081

seq documentation_block:

Eukaryota;

Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Euteleostomi; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 3128)
 Beneke, S., Meyer, R. and Buerkle, A.
 Isolation of cDNA encoding full-length rat (*Rattus norvegicus*) poly
 (ADP-ribose) polymerase
 Biochem. Mol. Biol. Int. 43 (4), 755-761 (1997)
 98046546
 2 (bases 1 to 3128)
 Beneke, S., Meyer, R. and Buerkle, A.
 Direct Submission
 Submitted (18-MAR-1997) Angewandte Tumoriologie, Deutsches
 Krebsforschungszentrum, Im Neuenheimer Feld 242, Heidelberg
 D-69120, Germany
 3 (bases 1 to 3128)
 Beneke, S., Meyer, R. and Buerkle, A.
 Direct Submission
 Submitted (19-FEB-1998) Angewandte Tumoriologie, Deutsches
 Krebsforschungszentrum, Im Neuenheimer Feld 242, Heidelberg
 D-69120, Germany
 Sequence update by submitter
 On Feb 19, 1998 this sequence version replaced gi:2583143.
 Location/Qualifiers
 1..3128
 /organism="Rattus norvegicus"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /cell_type="mononuclear blood cell"
 31..3075
 /codon_start=1
 /product="poly(ADP-ribose) polymerase"
 /protein_id="AAC53544.1"
 /db_xref="GI:2896792"
 /translation="MAEATERYRVEYAKSGRASCCKSESIPKDSLRMAIMVQSNF
 DGKVPWHVHSCFVGHSTIRPDTEVDQFSELRWDDQKVKTAAGGVAGQHGQ
 GGKAEKTLGDFAEYAKSNSTCKGMEKIEKGMRSLKAMLDPEKPLQMDIDRWHP
 TCFVKNDELGFRPEYASQLKGFSLLSAEDKELKQLPAVKSEKQDEVDGIDE
 VAKSKKDKDESSLEKALKQAQNELVNIKDELKACSTNDLKELLFQOQVPSG
 ESATLDVADGMAGFALPCQEGSLVFKSDAYCTGDTAWTKMVKTONPSREK
 VTPKREISYLKLIKQDLRFLPPSSAPAPPPVITISAPTAVNSAPADKPLS
 NMKLTILGKLSQNKDEAKMIEKLGKLTGSANKSLCISTKKEVEKSKMVEVKAA
 NRVVCEDFLDQYSASAKSLQELLSAHSLSHWAQVYKVEGVVFKGSAAPSKSKA
 GAVKEGVNKSERMKLTILKGAAYVDDSGLESHVLENGKGVFSATLGLVDIVKGT
 NSYKQLQSSDESRTWIFRSGRVGTIVGSKNLEQMPSEKEDAVHFMLKLEERTGN
 AMHSNKTYPKPYPLEIDYQDEEAQKLVKPGTKSLPKPQVQLGMIFDVESM
 KKALVEYIDLQKPLGKLSRRQIAAYSLSEVQAVSQSSQSLDLSNRFYELI
 PHDFGMKPPLLNNTDSVQAKVEMLDNLIDIEVAYSLRGGSDSKDPTDNYEKLK
 TDIKVDRDSEAEVIRKYVKNTHATTHAYDLEVIDIEKIEEGESQRYKPEQLHN
 RLLWHGSRRTNFAGILSGLRIAPPEAPVGYFGKGIYFADWMSKSNYCHTSQGD
 PIGLILLGEVALGNMYELKHASHISKLPKGVKGLGTAPDPSPASITLDGVEVPLG
 TGPSPGVNDTCLLYNETIVYDIAQVNLKLLKLFNFKTSLW"
 BASE COUNT 861 a 758 c 905 g 604 t
 ORIGIN

alignment_scores:
 Quality: 1592.00 Length: 1066
 Ratio: 2.331 Gaps: 36
 Percent Similarity: 64.071 Percent Identity: 36.961

alignment_block:
 US-09-236-995D-2 x RNU94340 ..

Align seg 1/1 to: RNU94340 from: 1 to: 3128

3 AlaProProLysAlaTrpLysAlaGluTyrAlaLysSerGlyArgAla 19
 |||
 40 GCCACGAGAGGCTTACCGAGTGGAGTACCTAAGACGGCGCGGCTC 89

19 rCysLysSerCysArgSerProLysAlaLysAspGlnLeuArgLeu 36
 |||
 90 TTGCAAGAAATGCCAGGAGAGTATCCCAAGAGACTCGCTCGCATGCCA 139

35 ysMetValGlnAlaSerGlnPheAspGlyPheMetProMetTrpAsnHis 52
 |||
 140 TCATGGTGCAGTCACCCATGTTTCGATGGAAAAAGTCCACACTGGTATCAC 189

53 AlaArgCysIlePheSerLysLysAsnGlnIleLysSerValasp...As 68
 |||
 190 TTCTCTCTGCTTCTGAAGGTGGCCACTCCATCCGGCAGCTTGACACAGA 239

68 pValGluGlyIleAspAlaLeuArgTrpAspGlnGluLysIleArgA 85
 |||
 240 GGTGACGGGTCTCTGAGTGGGTGATGACCCAGCAGAGGTCAAGA 289

85 snTyrVal.....GlySerAlaSerAlaGlyThrSerThrAla 98
 |||
 290 AGACCGCAGAGGCTGGAGGCTGGCAGCAAGCCAGCATGGAGGTGGC 339

99 AlaProProGluLysCys.....ThrIleGluIleAlaProSe 111
 |||
 340 GGCAGGCGGAGAGACCGCTGGCGACTTTGGCGCAGAGTATGCCAAGTC 389

111 rAlaArgThrSerCysArgCysSerGluLysIleThrLysGlySerV 128
 |||
 390 CAACAGGAGCAGTGCATGAGGCTGCATGAGAGATAGAGAGGTCAGA 439

128 aArgLeuSerAlaLys...LeuGluSerGluGlyProLys...GlyIle 142
 |||
 440 TCGGCTGTCCAAGAGATGCTGGATCCAGAGAGCCCACTGGGTATG 489

143ProTyrThrHisAlaAsnCysPhePheGlu..... 152
 |||
 490 ATTGACCGCTGGTACCATCCCACTTGTTCGCAAGAACCGGGACGAGCT 539

153ValSerProSerAlaThrValGluLysPheSerGlyTrpAspThrL 168
 |||
 540 GGGCTTCGGCTCAGTACAGTGCAGCCAGCTCAAGGGCTTTAGCGCTCC 589

168 euSerAspGluAspLysArgThrMet.....LeuAspLeuValLys 181
 |||
 590 TCTCTGCAGAGACAAAGAGCTCTGAAGAGAGCTCCCGCGCAGTCAAG 639

182 LysAspValGlyAsnAsnGluGlnAsnLysGlySerLysArgLysLys 198
 |||
 640 AGTGAA.....GGAAAGAGAAAAATGTGA 662

198 rGluAsnAspIleAspSerTyrLysSerAlaArgLeuAspGluSerThr 215
 |||
 663 CGAGCTGGAT.....GGAATAGATGAA..... 684

215 erGluGlyThrValArgAsnLysGlyGlnLeuValAspProArgGlySer 231
 |||
 685GTAGCCAAAAAGAAATCTAAA.....AAAGGAAAG 714

232 AsnThrSerSerAlaAspIleGlnLeuLysLysGluGlnSerAspTh 248
 |||
 715 GACAAGGAGAGTAGTAACCTGGAAAAAGCGCTCAAGGCTCAGAACGAGCT 764

248 rLeuTrpLysLeuLysAspGlyLeuLysThrHisValSerAlaAlaGlu 265
 |||
 765 GGTCTGGAATATCAACATCAGCTGAAGAAAGCTCTCTCACCACAGACC 814

265 euArgAspMetLeuGluAlaAsnGlyGlnAspThrSerGlyProGluArg 281
 |||
 815 TGAAGGAGTGTGCTCATCTTCAACAGCAGCAAGTACCATCTGGAGAGTCC 864

282 HisLeuLeuAspArgCysAlaAspGlyMetLeuPheGlyAlaLeuGlyPr 298
 |||
 865 GCATCTGACCGAGTGTGCTGATGCGGTGGCGCTTTGGGGGCCCTTCTGCC 914

298 oCysProValCysAlaAsnGlyMetTyrTyrAsnGlyGlnTyrGlnC 315
 |||
 915 CTCAAGGAGGTGTTTCAGGACAGCTGGTGTATTAAGAGCAGCGCTTATTACT 964

315 ysSerGlyAsnValSerGluTrpSerLysCysThrTyrSerAlaThrGlu 331

1823 GTAACAAGCTTGACGATGCCGTCCAAAGAGGATGCTGTGGAACACTTC 1872
581 LysArgLeuPheLeuGluLysThrGlyAsnSerTrpGluAlaTrpGluCy 597
1873 ATGAAGTTATACGAAGAAGACTGGAATGCCCTGGCACTCG 1914
597 sLysThrAsnPheArgLysGlnProGlyArgPheTrpProLeuAspVala 614
1915AAAAACTTCACAAAGTATCCCAAGAGTTCTACCCCTGGAGATTG 1960
614 sPyrGlyValLysLysAlaProLysArgLysAsp 625
1961 ACTATGGCCAGACGAAGAGCAGTAAAGAGCTGGCGTGAAGCCTGC 2010
626 lLeSerGluMetLysSerSerLeuAlaProGlnLeuLeuGluLeuMetLy 642
2011ACTAAGTCGAAGCTGCCGAAGCGGTGCAGGAGCTCGTCGG 2051
642 sMetLeuPheAsnValGluThrTyrArgAlaAlaMetMetGluPheGlu* 659
2052 GATGATCTTCGACGTGGAGCATGAAGAAGCCCTTGGTGAGTACGAGA 2101
659 ***AsnMetSerGluMetProLeuGlyLysLeuSer*****AsnIleGlu 675
2102 TTGACCTTCAGAAAGATGCCCTTGGGGAACCTGAGCAGAGGAGCAGATCCAG 2151
676 ***GlyPheGluAlaLeuThr*****LeuPheGluGlyHisAr 692
2152 GCGGCTACTCTATCTCAGTGAGGTCCAGCA 2183
692 g***SerSerThrGlyLeu***GluLysAla***LeuLeu***** 709
2184 GGCAGTGTCCCAAGGAGCAGTGAATCTCAGATCCTGGATCTCTCAATC 2233
709 *****PheSerLeuLeuSerLeuPheIleLeuIleLeuTyrGlyMet 725
2234 GCTTACACCCCTCATCCCCATGACTTTGGAATG 2268
726 ArgMetIleSerTyrSerLysAlaLys 734
2269 AAGAAGCCCTTCTGAACAACACAGACAGTGTGCAGGCCAAGGTGA 2318
734 sMetLeuGluAlaLeuGlnAspIleGluIleAlaSerLysIleVal 749
2319 GATGCTGGACACCTCTCGACATCAGGTGGCTATAGCTTCTTCAGGG 2368
750 ..GlyPheAspSerAspSerAspLeuSerLeuAspAspLysTyrMetLys 765
2369 GTGGGTCTGATGACAGCAGTAAGGACCCCATCGATGTCAACTATGAGAAG 2418
766 LeuHisCysAspIleThrProLeuAlaHisAspSerGluAspTyrLysLe 782
2419 CTCAAAACTGACATTAAGGTGGTTGACAGAGATTCTGAAGAGCTCAGGT 2468
782 uIleGluGlnTyrLeuLeuAsnThrHisAlaProThrHisLysAspTrps 799
2469 CATCAGGAAGTATGTGAAGAACACTCAGCGACCCACGACCAATGCCATG 2518
799 erLeuGluLeuGluValPheSerLeuAspArgAspGlyGluLeuAsn 815
2519 ACCTGGAAGTGATAGACATCTTTAAGATAGAGCGAGAGGAGAGCCAA 2568
816 LysTyrSerArgTyrLysAsnAsnLeuHisAsnLysMetLeuLeuTrpHi 832
2569 CGGTACAAGCCCTTCAGG...CAGCTTCAACACCGGAGACTGCTGTGGCA 2615
832 sGlySerArgLeuThrAsnPheValGlyIleLeuSerGlnClyLeuArgI 849
2616 CGGTCCAGGACCAACCAACTTCGAGGACATCTCTCACAGGCTCTGGGA 2665
849 leAlaProProGluAlaProValThrGlyTyrMetPheGlyLysGlyLeu 865
2666 TAGCCCCACCTGAAGACCTGTGACAGGCTACATGTTTGGGAAAGGAATC 2715

965 GTACTGGGATGTCACCTGCTGACCAAGTGCATGCTCAAGACACAGAAAT 1014
332 ProValArgValLysLysLysTrpGlnIleProHisGlyThrLysAsnAs 348
1015 CCAGCCGAAAGGAATGGTAACTCCAAAGGAATT 1049
348 pTyrLeuMetLysTrpPheLysSerGlnLysValLysLysProGluArgV 365
1050 CCAGAAATATCTACTCTCAAGAAATTAAGATCAAAAGACGAGCCGCG 1099
365 allLeuProMetSerProGluLysSerGlySer 376
1100 TGTTCCTCCAGAGACAGCGCCCGCCAGCACCACCACTCCCTCCGCTCG 1149
377 LysAlaThrGlnArgThrSerLeuLeuSerSerLysGlyLeuAspLys... 392
1150 ATCACCCTCCGACCCACAGCTGTGAAGTCTCTCTGACACAGCAGCAAAAC 1199
393LeuArgPheSerValValGlyGlnSerLysG 403
1200 CTTGCTTAACATGAAGATCCTGACTCTTGTGAAGCTCTCCCAAGCAAGG 1249
403 luAlaAlaAsnGluTrpIleGluLysLeuLysLeuAlaGlyAla 417
1250 ACGAAGCAAGGCATGATTGACAAACTCGAGGCAAGCTGACAGGATCT 1299
418 AsnPheTyrAlaArgValValLysAspIleAspCysLeuIleAlaCysGl 434
1300 GCCAACAAGGCTCTCTTGTGTATCAGCACTAAGAA 1334
434 yGluLeuAspAsnGluAsnAlaGluAlaArgLysAlaArgLeuLysI 451
1335 GGAGGTGGAGAAGATGATGAAGATGAGGAAGTGAAGTGAAGGAGCAGCCAA 1384
451 leProIleValArgGluGlyTyrIleGlyGluCysValLysArgThrLys 467
1385 TTCAGTGTGTGTGAGGACTCTCTCAGGATGTCTGCTGCTCAGCTAAG 1434
468CysCysHisLeuIleCysIleAsnTrpAsnAl 478
1435 AGCTCCCAAGAGCTGCTCTCAGCCCATAGCTTGTC...TCGTGGGGGC 1481
478 a...LeuGluSerSerLysGly***ThrValThrValLysValLysGlyA 494
1482 TGAGGTGAAGTAGAGCTGTGAAGTGGTGGTCCCCAAGGGAGGATCAG 1531
494 rgSerAlaCysSer***SerPro***ValCysLysAsnThrAlaHisIle 510
1532 CTGCACCTCCCAAGAGAGCAAGGTGTGTCAAGGAGGAAGT 1575
511 Pro*****TrpGluLysHisIleGlnCys***LeuLys 523
1576 GTCACAAAGTCTGAAGAGAGGATGAATAACTAAGCTGTAAGGAGGAGGAGC 1625
524HisValLeu 526
1626 CTTGATCCTGACTGTGCTGGACACTCTGCACATGCTCTGGAGAGAG 1675
527Thr***His***Val 531
1676 GCGGAAGGTGTTACGTGCCAGCTGGTGGTGGACATTTGTGAAGGG 1725
532 CysThrGlyTyrTyrValLeuGlnIleIleGluGlnAspAspGlySerGl 548
1726 ACTAATCTCTATTACAACTGCAACTCTGGAGTCCGACAGGAGGAGCAG 1775
548 ucYstYrValPheArgLysTrpGlyArgValGlySerGluLysIleGlyG 565
1776 GTATTGGATCTTCGGCTCTCGGGCCGAGTGGGCACA...GTATCGGCA 1822
565 lylGlnLysLeuGluMet...SerLysThrGluAlaIleLysGluPhe 580
.....

866 TyrPheAlaAspLeuValSerLysSerAlaGlnTyrCysTyrValAspAr 882
 |||||
 2716 TACTTTGCTGATATGGTGTCCAAAGTCCGAACACTACTCCACAGCTCTCA 2765
 882 gAsnAsnProValGlyLeuMetLeuLeuSerGluValAlaLeuGlyAspM 899
 |||||
 2766 GGGAGACCCGATGGCTTAACTGTTGGGAGAAAGTTGCCCTTGGAAACA 2815
 899 eTyrGluLeuLysLysAlaThrSerMetAspLysProProArgGlyLys 915
 |||||
 2816 TGTACGAACCTCAAGCATCTCTCACATCAGCAAGTTACCCAAAGGCAAG 2865
 916 HisSerThrLysGlyLeuGlyLysThrValProLeuGluSerGluPheVa 932
 |||||
 2866 CACAGTGTCAAGAGTTTGGGCAAAACCGCCCTGACCCCTTCGGCCAGCAT 2915
 932 lLysTrpArgAspValValProCysGlyLysProValProSerS 949
 |||||
 2916 CACC...CTGGATGCTGAGAGTTCCGCTGGGAACAGGGATTCCGCTG 2962
 949 eRIleArgSerSerGluLeuMetTyrAsnGluTyrIleValTyrAsnThr 965
 |||||
 2963 GTGTTAATGACACCTGCTGCTGTATACCGAGTACATTTGCTACGACATT 3012
 966 SerGlnValLysMetGlnPheLeuLeuLysValArgPheHisLys 981
 |||||
 3013 GCTCAGGTGAATCTGAAGTACCTACTGAAACTGAAGTTCAATTCAAG 3060

seq_name: gb_pat:114359

seq_documentation_block: 3747 bp DNA PAT 26-SEP-1995
 LOCUS 114359
 DEFINITION Sequence 5 from patent US 5449605.
 ACCESSION 114359
 VERSION 114359.1 GI:996850

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

METHOD OF DETECTING A PREDISPOSITION TO CANCER BY DETECTING A

DELETION POLYMORPHISM IN THE GENE FOR HUMAN POLY (ADP-RIBOSE)

POLYMERASE

POLYMERASE

PATENT: US 5449605-A 5 12-SEP-1995;

LOCATION/QUALIFIERS

1. 3747

/organism="unknown"

BASE COUNT 1036 a 844 c 1060 g 807 t

ORIGIN

alignment_scores:

Quality: 1587.00 Length: 1060

Ratio: 2.290 Gaps: 36

Percent Similarity: 65.377 Percent Identity: 36.226

alignment_block:

US-09-236-995D-2 x 114359 ..

Align seg 1/1 to: 114359 from: 1 to: 3747

6 LysAlaTrpLysAlaGluTyrAlaLysSerGlyArgAlaSerCysLysSe 22

|||

209 AAGCTCTATCGAGTCAGTACGCAAGACGGCGCGCTCTTGAAGAA 258

22 rCysArgSerProIleAlaLysAspGlnLeuArgLeuGlyLysMetValG 39

|||

259 ATGCACGAGACATCCCAAGAGACTCGCTCGGATGGCCATCATGGTGC 308

|||

39 lAlaSerGlnPheAspGlyPheMetProMetTyrAsnHisAlaArgCys 55

|||

|||

|||

|||

|||

|||

|||

|||

|||

309 AGTCGCCCATGTTTATGGGAAAGTCCACACTGGTACCACATCTCTCTCTGC 358
 56 IlePheSerLysLysAsnGlnIleLysSerValAsp...AspValGluG1 71
 |||||
 359 TTTCTGGAAGGTGGGCCACTCTCATCCGCCACCCCTGACGTTGAGGTGGATGG 408
 71 yIleAspAlaLeuArgTrpAspAspGlnGluLysIleArgAsnTyrVal. 87
 |||||
 409 GTTCTCTGACCTTCGGTGGATGACCCAGAGAAAGTCAAGAACAGACAGCG 458
 88GlySerAlaSerAlaGlyThrSerSerThrAlaAlaProPro 101
 |||||
 459 AAGCTGGAGGAGTGCACAGCAAAAGGCCAGGATGGAATTGGTAGCAAGGCA 508
 102 GluLysCys.....ThrIleGluIleAlaProSerSerAlaArgTh 114
 |||||
 509 GAGAAAGACTCTGGGTGACTTTTCAGCAGAGAGTATGCCAAGTCCCAACAGAAG 558
 114 rSerCysArgArgCysSerGluLysIleThrLysGlySerValArgLeuS 131
 |||||
 559 TAGGTCCAAGGGGTGTATGAGAGAGATAGAAAAGGCCAGGTGCCCTCTG 608
 131 eAlaLysLeu...GluSerGluGlyProLys...GlyIle.....Pro 143
 |||||
 609 CCAAGAAGATGTTGGACCCGAGAGACACAGCTAGGCATGATTGACCGC 658
 144 TTPTrpHisAlaAsnCysPhePheGlu.....ValSe 154
 |||||
 659 TGGTACCATCCAGGCTGCTTTCTCAAGAACAGAGGAGGAGCTGGGTTTCCG 708
 154 rProSerAlaThrValGluLysPheSerGlyTrpAspThrLeuSerAspG 171
 |||||
 709 GCCCGAGTACAGTGGCGATCAGCTCAAGGGCTTCAGCCTCTCTGTGTACAG 758
 171 luAspLysArgThrMetLeuAspLeuValLysLysAspValGlyAsnAsn 187
 |||||
 759 AGGATAAA.....GAAGCCCTGAAGAAGACAGCTCCAGGAGTC 796
 188 GluGlnAsnLysGlySerLysArgLysLysSerGluAsnAspIleAspSe 204
 |||||
 797 AAG.....AGTGAAGGAAGAAAGAAAGCGGATGAGGTGGATGG 834
 204 rTyrLysSerAlaArgLeuAspGluSerThrSerGluGlyThrValArg 221
 |||||
 835 A.....GTGGATGAAGTGGCGAAGAAAGAAATCTAAAG 869
 221 snLysGlyGlnLeuValAspProArgGlySerAsnThrSerSerAlaAsp 237
 |||||
 870 AAAAAGACAAG.....GATAGTAAG 889
 238 IleGlnLeuLysLysGluGlnSerAspThrLeuTrpLysLeuLysAs 254
 |||||
 890 CTTGAAAAGCCCTAAGGCTCAGACAGCAGCTGATCTGGAAACATCAAGGA 939
 254 pGlyLeuLysThrHisValSerAlaGluLeuArgAspMetLeuGluA 271
 |||||
 940 CGAGCTAAAGAAAGTGTCTCACTAATGACCTGAGGAGCTACTCATCT 989
 271 laAsnGlyGlnAspThrSerGlyProGluArgHisLeuLeuAspArgCys 287
 |||||
 990 TCAACAAGACAGCAAGTCTCTGGGAGTGGCGGATCTTTGGACCCAGTA 1039
 288 AlaAspGlyMetLeuPheGlyAlaLeuGlyProCysProValCysAlaAs 304
 |||||
 1040 GCTGATGCGATGGTTCGGTGGCTCCTCTCTCCCTGGCAGGAATGCTCGG 1089
 304 nGlyMetTyrTyrAsnGlyGlnTyrGlnCysSerGlyAsnValSerG 321
 |||||
 1090 TCAGCTGGTCTTCAAGAGCAGTGCCTATTACTGCACCTGGGAGCTCACTG 1139
 321 luTrpSerLysCysThrTyrSerAlaThrGluProValArgValLysLys 337
 |||||
 1140 CTTGGACCAAGTGTATGGTCAAGACACAGACACCAACCGG.....AAG 1183

584 PheLeuGluLysThrGlyAsnSerTrpGluAlaTrpGluCysLysThrAs 600
... ||||| ||||| ||||| ||||| ||||| :
2042 TATGAGAAAAACCGGAACGCTTGGCACTCC AAAA 2079

600 nPheArgLysGlnProGlyArgPheTyrProLeuAspValAspTyrGlyV 617
|||| | ||| : ||||| ||||| ||||| ||||| :
2080 TTTCACCAAGATATCCCAAAGTTTACCCTCGGAGATTGACTATGGCC 2129

617 al... LysLysAlaProLysArgLysAspIleSerGlu...MetLysSer 631
||||| : ||||| ||||| :
2130 AGATGAAGAGGCAGTCAAGAAGCTCACAGTAATCTCTGGCACAAGTCC 2179

632 SerLeuAlaProGlnLeuLeuLeuMetLysMetLeuPheAsnValG 648
:||| : ||||| ||||| ||||| ||||| :
2180 AAGTCTCCCACAGCAGTTTCAGGACCTCATCAAGATGATCTTTGATGGGA 2229

648 uThrTyrArgAlaAlaMetMetGluPheGlu***AsnMetSerGluMetP 665
:||| : ||||| ||||| ||||| ||||| :
2230 AAGTATGAAGAAAGCCATGTGTGAGTATGAGATCGACCTTCAGAAGATGC 2279

665 roLeuGlyLysLeuSer*****AsnIleGlu**GlyPheGluAlaLeu 681
:||| : ||||| ||||| ||||| ||||| :
2280 CTTGGGCAAGCTGAGCAAAAGCAGATCCAGGCCGATACCTACTCCATCCTC 2329

682 Thr*****LeuPheGluGlyHisArg***SerSerThrGlyLe 698
:||| : ||||| ||||| ||||| :
2330 AGT GAGTCCACAGCGGTGTCTCAGGCGAG 2361

698 u**GluLysAla**LeuLeu*****PheSerLeuLeuS 715
:||| : ||||| ||||| :
2362 CAGGACTCTCAGATCCTGCATCTCTCAATCGCTTTTACACCTTGATGCC 2411

715 erLeuLeuPheIleLeuTyrGlyMetArgMet..... 727
:||| : ||||| ||||| :
2412 CCCAGCACTTT GGGATGAAGAGCCCTCCGCTCCTG 2446

728 IleSerTyrSerLysAlaLysMetLeuGluAlaLeuGl 740
:||| : ||||| ||||| :
2447 ACATATGCAGACAGTGTGAGGCCAAGGTGGAATGCTTGACAACCTGCT 2496

740 nAspIleGluIleAlaSerLysIleVal..... GlyPheAspSerAsps 755
:||| : ||||| ||||| :
2497 GGACATCAGGTGGCTTACAGTCTGCTCAGGGAGGCTCTGATGATGACA 2546

755 erAspGluSerLeuAspLysTyrMetLysLeuHisCysAspIleThr 771
:||| : ||||| ||||| :
2547 GCAAGATCCCATCGATGCAACTATGAGAAGCTCAAACTGACATTAAG 2596

772 ProLeuAlaHisAspSerGluAspTyrLysLeuIleGluGlnTyrLeu 788
:||| : ||||| ||||| :
2597 GTGTGTACAGAGATTCTGAAGAAGCCGAGATCATCAGGAAGTATGTTAA 2646

788 uasnThrHisAlaProThrHisLysAspTrpSerLeuGluLeuGluLuv 805
:||| : ||||| ||||| :
2647 GAACACTCATGCACACACACAGTGTGATGACTTGGAAAGTCATCGATA 2696

805 alPheSerLeuAspArgaspGlyGluLeuasnLysTyrSerArgTyrLys 821
:||| : ||||| ||||| :
2697 TCITTAAGTAGAGGTGAAGCGAATGCCAGCTTTACAAGCCCTTTAAG 2746

822 AsnAsnLeuHisAsnLysMetLeuLeuThrHisGlySerArgLeuThrAs 838
:||| : ||||| ||||| :
2747 ...CAGCTTCATAACCGAAGATGCTGTGGCACGGGTCCAGGACCACCAA 2793

838 nPheValGlyIleLeuSerGlnGlyLeuArgIleAlaProGluAlaAp 855
:||| : ||||| ||||| :
2794 CTITGTGGGATCCTCTCCAGGGTCTTCGGATAGCCCCCTCGAAGCGC 2843

855 roValThrGlyTyrMetPheGlyLysGlyLeuTyrPheAlaAspLeuVal 871
:||| : ||||| ||||| :
2844 CCCTGACAGGCTACATGTTTTGGTAAAGGGATCTATTTCGCTGACATGCTC 2893

872 SerLysSerAlaGlnTyrCystyrValAspArgAsnProValGlyLe 888

```
2894 TCACAGAGTGCACCACTACCTACCTCAGGAGACCAATAGGCTT 2943
888 uMetLeuSerGluValAlaLeuGlyAspMetTyrGluLeuLysLysA 905
2944 AATCCTGTTGGGAGAGTTGGCTTGGAAACATGTATGAACAGGACG 2993
905 laThrSerMetAspLysProProArgGlyLysHisSerThrLysGlyLeu 921
2994 CTTACATATACAGCATGAGTACCAAGGCAAGCAGTGTCAAAGTTG 3043
922 GlyLysThrValProLeuGluSerCluPheValLysTrpArgAspVa 938
3044 GCACAACTACCTGATCTTACCTAACATTAGT...CTGGATGGTGT 3090
938 lValValProCysGlyLysProValProSerSerIleArgSerSerLul 955
3091 AGACGCTTCTCTGGGCGGATTCATCTGCTGATGTAATGACACCTCT 3140
955 euMetTyrAsnGluTyrIleValTyrAsnThrSerGlnValLysMetGln 971
3141 TACTATATAAGCAGTACATTTCTATGATATGCTCAGTAAATCTGAAG 3190
972 PheLeuLeuLysValArgPheHisHisLys 981
3191 TATCTGCTGAACTGAATCAATTTTAAG 3220
seq_name: gb_ro:MMADPRP
seq_documentation_block:
LOCUS MMADPRP 3172 bp mRNA ROD 12-SEP-1993
DEFINITION Mouse mRNA for poly (ADP-ribose) polymerase (EC 2.4.2.30).
ACCESSION X14206
VERSION X14206.1 GI:49893
KEYWORDS DNA-binding protein; NAD(+) ADP-ribosyltransferase; zinc finger protein.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 3172)
Huppi,K.
Direct Submission
Submitted (26-JAN-1989) Huppi K., NCI/NIH, Lab of Genetics, Building 37, Room 2B-21, Bethesda, MD 20892
2 (bases 1 to 3172)
Evers,R., Hammer,A. and Cornelissen,A.W.C.A.
Unusual C-terminal domain of the largest subunit of RNA polymerase II of Chithidia fasciculata
Nucleic Acids Res. 17, 3387-3401 (1989)
Location/Qualifiers
1..3172
/organism="Mus musculus"
/strain="BxSB"
/db_xref="taxon:10090"
/tissue_type="spleen"
/clone="pADP3-2.1 and pADP-3C"
85..3126
/notes="poly (ADP-ribose) polymerase (AA 1 - 1013)"
/codon_start=1
/protein_id="CAA32421.1"
/db_xref="GI:49894"
/db_xref="MGI:87933"
/db_xref="SWISS-PROT:P11103"
/translation="NAEASRLRYQYAKSGRACKKSESIPKDSLRMAIMVQSPMF
DGKVPWHYHFCFVKVGSIRHPDVEGFSRLRWDQKVKKTAEGAGVAGRGQDS
GGKAEKTLGDPAEAKNSRMCCKLEIKGQMRISKWMDPEKPLQGMIDRWHP
TCFVKRDELGRPYBSASQLKSLSAEDKALKQLPAIKNEGRKRGDEVDTPE
VAKKSRDETDKRYLKAQNELIWNIDELKACSTNDKELIFNQOQVPSGE
SAILDRADGMAFGALLPCCKSGOLVFKDYYCTGDTAVTAKCMVKTONPBRKEW
TPKEFETSYLKLKVKQDRIFPESSAPITVHPLSVTSAPAVNSAPADPKLPSN
MKILTLGLSONKDEKAVIEKLGKLTGSANKASLCISIKKEVERNNKMEVKKAN
IRVSEDFLDQV/SASTKSLQDLLSAHLSLSPWGAEPVEFVAPRGKSAAPSKKSG
```

```
CFKEGVNKSSEKRMKLTLLKGAADVDPDSGLEHSAHVLEKGGKVSATLGLVDIVKGTN
SYKLOLLEDDKESRYWIFRSWGRCLCTVIGSNKLBOMPSKEEAVQPMKLYEETGNA
WHSKNFTKPKFYPLEIDYGOEBAVKLTVPKGTSKLPKQVQLVGMIFEDVDSMK
KALVEYEIDLOKMPGLKSRROIQAAVSTLSEVQOPVQSGSSESOILDLNRFYTLIP
HDFGMDKPPLLNADSVQAKVEMLDNLIDIEVAYSLRGGSDSDSKDPIDVNYEKLKT
DIKVPDRDSEEAIEVIRKYKYNTHATTHAYDLVIEDIFKIEREGESQRYKPFRLQHNH
RLLMHGRITTFAGLISQGLRAPPAPVPTGYMFGKGIYFADWVSKSANYCHTSGQDP
IGLIMIGEVALGNMYELHSHISKPKGHSVGLGKTTDPDPSASITLEGVEVPLGT
GIPSGVNDTALLYNEYIVYDTAQVNLKYLKLPKFKTSLW"
88..678
misc_feature /note="zinc finger domains"
2743..2823
misc_feature /note="DNA-binding domain A"
2898..2952
misc_feature /note="DNA-binding domain B"
BASE COUNT 869 a 779 c 925 g 599 t
ORIGIN
alignment_scores:
Quality: 1584.50 Length: 1081
Ratio: 2.340 Gaps: 37
Percent similarity: 62.627 Percent Identity: 36.078
alignment_block:
US-09-236-995D-2 x MMADPRP
Align seg 1/1 to: MMADPRP from: 1 to: 3172
3 AlaProProLysAlaTriLysAlaGluTyrAlaLysSerGlyArgAlase 19
|||
94 CCCTCGGAGAGGCTTTTCGAGTGCAGTACGCAAGCGGCGCGCTC 143
19 GCCTLysSerCysArgSerProIleAlaLysAspGlnLeuArgLeuGlyL 36
|||
144 TTGCAAGAAATGCGAGCAGAGTATTCCTCCAAAGCACTCCCTCCGATGGCCA 193
36 YsMetValGlnAlaSerGlnPheAspGlyPheMetPrometTirAsnHis 52
|||
194 TCATGGTGGAGTCCCATTTTCGATGGGAAAGTCCACACTGGTATACCAC 243
53 AlaArgCysIlePheSerLysLysAsnGlnIleLysSerValAsp...As 68
|||
244 TTCTCTCTGCTTCTGGAAGTGGCGCAGTCCATCCGCGCACCTGATGTGA 293
68 pValGluGlyIleAspAlaLeuArgTrpAspGlnGluLysIleArgA 85
|||
294 GGTGATGGCTTCTCTGAGTCCGCTGGGATGATCAGCAGAAAGTCAAGA 343
85 snTyrVal.....GlySerAlaSerAlaGlyThrSerSerThrAla 98
|||
344 AGACGGCCGAGCTGGAGCGCTGGCAGCAAGCGCAGGATGGAGGTGG 393
99 AlaProProGluLysCys.....ThrIleGluIleAlaProSe 111
|||
394 GGCAAGGGCGGAGAGACATTTGGTGACTTTTGCAGCAGATGATGCCAAGTC 443
111 rAlaArgThrSerCysArgCysSerGluLysIleThrLysGlySerV 128
|||
444 CAACAGGAGCATGTGCAAGGCTGCTGGAGAGATAGAGAAGGGCCAGA 493
128 alArgLeuSerAlaLysLeu...GluSerGluGlyProLys...GlyIle 142
|||
494 TGCCTCTGCTCCAGAAAGATGGTGATCCAGAGAAAGCCACAGCTGGGTATG 543
143 .....ProTyrTyrHisAlaAsnCysPhePheGlu..... 152
|||
544 ATTGACCGCTGGTACCATTCCAACTTCCTTTGTCAAGAAAGCGGACGAGCT 593
153 ....ValSerProSerAlaThrValGluLysPheSerGlyTirAspThrL 168
|||
594 GGGCTTCGCGCGCAGTACAGTCCAGCTCAGCTCAGGCTTTAGCCTCC 643
```

[illegible]

2240 TGCTCAAGCAGCAGTGATCCAGATCTTAGATCTCTCCAATCGTTC 2280

711 PheSerLeuLeuSerLeuPheIleLeuIleLeuTyrGlyMetArgMe 727
2290 TACACTCTGATCCCCCATGACTTT.....GGATGAAGA 2324
727 t.....IleSerTyrSerLysAlaLysMetL 736
2325 GCOCCTCACTCCTGAACAACGCACAGCGTGACAGGCCAAGGTGGAGATGC 2374

736 euGluAlaLeuGlnAsnIleGluIleAlaSerLysIleVal.....Gly 750
2375 TAGACAACCTCTCTGGACATCGAGGTGGCCTATAGTCTCTTCAGGGGTGGC 2424

751 PheAspSerAspSerAspGluSerLeuAspAspLysTyrMetLysLeuH1 767
2425 TCTGAGACAGCAGCAGGAGTCCATCGACGTCAACTACCAGAAATCAA 2474

767 sCysAspIleThrProLeuAlaHisAspSerGluAspTyrLysLeuIleG 784
2475 AACTGCATTAAGTGTGTACAGAGATTCTGAAGAGCGCGAGGTGCATCA 2524

784 luGlnTyrLeuLeuAsnThrHisAlaProThrHisLysAspTrpSerLeu 800
2525 GGAAGTAGCTGAAGAACACATCATGCTACCACGCACAAACGCTATGACCTG 2574

801 GluLeuGluGluValPheSerLeuAspArgAspGlyGluLeuAsnLysTy 817
2575 GAAGTGATCGATATCTTCAAGATAGACGCGAGGGGAGCGCAGCGCTA 2624

817 rSerArgTyrLysAsnAsnLeuHisAsnLysMetLeuLeuTrpHisGlys 834
2625 CAAGCCCTTCAGG...CACGTTACAACCGAGGTGCTGTGCACGGCT 2671

834 erArgLeuThrAsnPheValGlyIleLeuSerGlnGlyLeuArgIleAla 850
2672 CCAGGACCAACCAACTTGTGTGCATCCTGTGCAGGGTCTGCGGATAGCC 2721

851 ProProGluAlaProValThrGlyTyrMetPheGlyLysGlyLeuTyrPh 867
2722 CCACCTGAAGCCCGCTGACAGCTACATGTTGGAAAGGATCTACPT 2771

867 eaLaAspLeuValSerLysSerAlaGlnTyrCystyrValAspArgAsnA 884
2772 TGCCGACATGTTGTCCAAAAGTGCAAACTACTGTCACACATCTCAGGGAG 2821

884 snProValGlyLeuMetLeuSerGluValAlaLeuGlyAspMetTyr 900
2822 ACCCGATTGGCTTAATAATGCTGGGAGAGTTGCCCTTGGAAATATGAT 2871

901 GluLeuLysLysAlaThrSerMetAspLysProProArgGlyLysHisSe 917
2872 GAACTCAAGCATGCTTCATATATCAGCAAGTTACCCAGGGCAACACAG 2921

917 rThrLysGlyLeuGlyLysThrValProLeuGluSerGluPheValLysT 934
2922 TGTCAAGAGTTTGGAAAAACACCCCTGACCCCTTGGCCGACGATCACCC 2970

934 rpArgAspAspValValValProCysGlyLysLysProValProSerSeri 950
2971 ..CTGAGGGGTAGAGGTTCCACTGGGAACAGGATCCCATCTGGGTGC 3018

951 ArgSerSerGluLeuMetTyrAsnGluTyrIleValTyrAsnThrSergL 967
3019 AACGACACTGCCCTCTGCTATATAGTACATCTGCTACGACATGCTCA 3068

967 nValLysMetGlnPheLeuLeuLysValArgPheHisLys 981
3069 GGTGAATCTCAAAATACCTGCTGAACATCAAGTTCAATTTTAAG 3111

seq_name: gb_pat:AX058340
seq_documentation_block:

seq_name: qb_pat:AX058340

seq_documentation_block:

LOCUS	AX058340	3045 bp	DNA	PAT	17-JAN-2001
DEFINITION	Sequence 24 from Patent WO00771179.				
ACCESSION	AX058340				
VERSION	AX058340.1 GI:12310800				
KEYWORDS	human..				
SOURCE	ORGANISM Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 3045)				
AUTHORS	Christenson,E., Demaggio,A.J., Goldman,P.S. and Mcelligott,D.L.				
TITLE	Human poly(adp-ribose) polymerase 2 materials and methods				
JOURNAL	Patent: WO 0071179-A 24 21-DEC-2000;				

FEATURES
SOURCE

CDS

```
1. .3043
/note="unnamed protein product"
/codon_start=1
```

[illegible]

alignment scores:

Quality:	1584.00	Length:	1060
Ratio:	2.289	Gaps:	36
Percent Similarity:	65.283	Percent Identity:	36.226

alignment_block:

US-09-236-995D-2 x AX058340

Align seg 1/1 to: AX058340 from: 1 to: 3045

6 LysAlaTrpLysAlaGluTyrAlaLysSerGlyArgAlaSerCysLys 22
||| : : : : :
||| : : : : :

19 AAGCTCTATCGAGTCGAGTACGCCAAGAGCGGGCGCGCTCTTTCGAAGAA 68

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1

22 rCysArgSerProIleAlaLysAspGlnLeuArgLeuGlyLysMetValG 39

69 ATGCAGGAGAGCATCCCCAAGGACTCGTCCGGATGGCCATCATGGTGC 118

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

39 InAlaSerGlnPheAspGlyPheMetProMetTrpAsnHisAlaArgCys 55

119 AGTCGCCCATGTTTGATGGAAAAGTCCACACTGGTACCACTTCTCCTGC 168

56 T l o p h o s o r i v o i n a n o c i n t i o n g o u l e s e s

```

56 IlePheSerLysLysAsnGlnFileLysSerValAsp...AspValGluGl 71
   :..... :.....| | | | | .....| | | | |

```

169 TTCTCGAAGGTGGGCCCACTCCATCCGGCAACCTGACGTTGAGGTGGATCC

103 TACGGAAAGGAGGGCCACATCCGACCCCTGACGTTGAGGTGGATGG Z118

71 yIleAspAlaLeuArgTrpAspAspGlnGluLysIleArgAsnTyrVal. 87

[illegible]

alignment_block:

US-09-236-995D-2 x AX062277

Align seq 1/1 to: AX062277 from: 1 to: 3045

6 LysAlaThrPlySalGluTyrAlaLysSerGlyArgAlaSerCysLysSse 22
||| :::: ||::||::||::||::||::||::||::||::||::||::||
19 AAGCTCTATCGAGTCGATACGCCAAGAAGGGCCGCCTCTTTGCAAGAA 68

22 rCySarGerProIleAlaLysAspGlnLeuArgLeuGlyLysMetValc 39
:||| :::: ||| ||||||||::||::||::||::||::||::||::||
69 ATGCAGCAGACATCCCCAAGACTCGCTCCGGATGGCCATCATGTGC 118

39 InAlaSerGlnPheAspGlyPheMetProMetTrpAsnHisAlaArgCys 55
||::: ::||::||::||::||::||::||::||::||::||::||
119 AGTCGCCCATGTTTGATGGAAGATCCCACACTGGTACCACCTCTCTCTGC 168

56 IlePheSerLysLysAsnClnIleLysSerValasp... AspValGlul 71
TTTT:::: ::::||::||::||::||::||::||::||::||::||
169 TTTCTGGAAGTGGGCCACTTCATCCGCCACCCCTGACCTGAGGTGGATGG 218

71 yIleAspAlaLeuArgTTPAspAspGlnGluLysIleArgAsnTyrrVal. 87
||||: ||||||||::||::||::||::||::||::||::||::||:
219 GTTCTCTGAGCTTCGGTGGGATGACCAGCAGAAAGTCAAACAACACGGC 268

88 GlySerAlaSerAlaGlyThrSerSerThrAlaAlaProPro 101
||| :::: ||| ::::||::||::||::||::||::||::||::||:
269 AAGCTGGAGGAGTAGACGGCAAGGCCAGGATGGAATTGTGTAGCAAGGCA 318

102 GluLysCys..... ThrIleuIleAlaProSerAlaArghH 114
||||| :::: ||| |||||::||::||::||::||::||::||::||
319 GAGAACAGCTCGGTGACTTTGACAGCAGAGPATGCCAAGTCCAAACAGAG 368

114 rSerCysArgArgCysSerGluLysIleThrLysGlySerValargLeus 131
:::||||: ||| ||||||||::||::||::||::||::||::||::||
369 TAGCTGCAAGGGTGTATGGAAGATAGAAAAGGGCCAGGTGGCCTGT 418

131 exAlaLysLeu... GluSerGluGlyProlys... Glylle..... Pro 143
|| |::: ||: || |::: ||::||::||::||::||::||::||:
419 CCAAGAAGATGTGACCGCGAGGAAGCACAGCTAGCATGATTGACCGC 468

144 TpTYrHisAlaAsnCysPhePheGlu..... Valse 154
|||::||::||::||::||::||::||::||::||::||::||:
469 TGTAACATCCAGGTGCTTGTGTCAGNACAGGAGGAGCTGGTTCG 518

154 rProSerAlaThrValGluLysPheSerGlyTTPAspThrLeuSerAspg 171
||::: ::::||::||::||::||::||::||::||::||::||::||
519 GCCCGAGTACATGCGAGTCAGCTCAAGGGCTTCAGCTCTCTGCTACAG 568

171 luAsplyArgthrMetLeuaspleuVallLysLysAspvalGlyAsnAsn 187
||||||| :::: ||: ||||||||::||::||::||::||::||::||
569 AGGATAAA..... GAAGCCCTGGAAGAGCAGTCCCGAGGAGTC 606

188 GluClnAsnLysCylserLyArgLysSerGluAsnAspileaspse 204
::: ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||:::
607 AAG..... AGTCAGGAAGAAAGAAAGCGCATGAGGTGGATGG 644

204 rTyrLysSerAlaArgLeuaspluSerThrSerGluClYThrValarga 221
||||||| :::: ||::||::||::||::||::||::||::||::||
645 A..... GTGGATGAAGTCGGCGAAGAAAGAAATCTAAAAAAG 679

221 snLysGlyGlnLeuValaspProArgGlySerAsnThrSerSerAlaasp 237
:||| :::: ||| ::::||::||::||::||::||::||::||::||:
580 AAAACACAAAG..... GATAGTAAG 699

238 IleclnLeuLysLeuLysGluClnSerAspThrLeutrpLysLeuLysAs 254
||||| ||::||::||::||::||::||::||::||::||::||
700 CTTGAAAAAGCCCTAAAGGCTCAGAACACACCTGTCTGGAAACATCAAGGA 749

254 pGlyLeuLysThrHisValSerAlaAlaGluLeuArgaspmetLeuclu 271
||| ||::||::||::||::||::||::||::||::||::||
750 CGAGCTAAAAAAGTGTGTCAACTTAATGACCTGAAGGAGCTACTCATCT 799

271 laAsnGlyGlnAspThrSerGlyProGluArgHisLeuLeuasparGcys 287

```

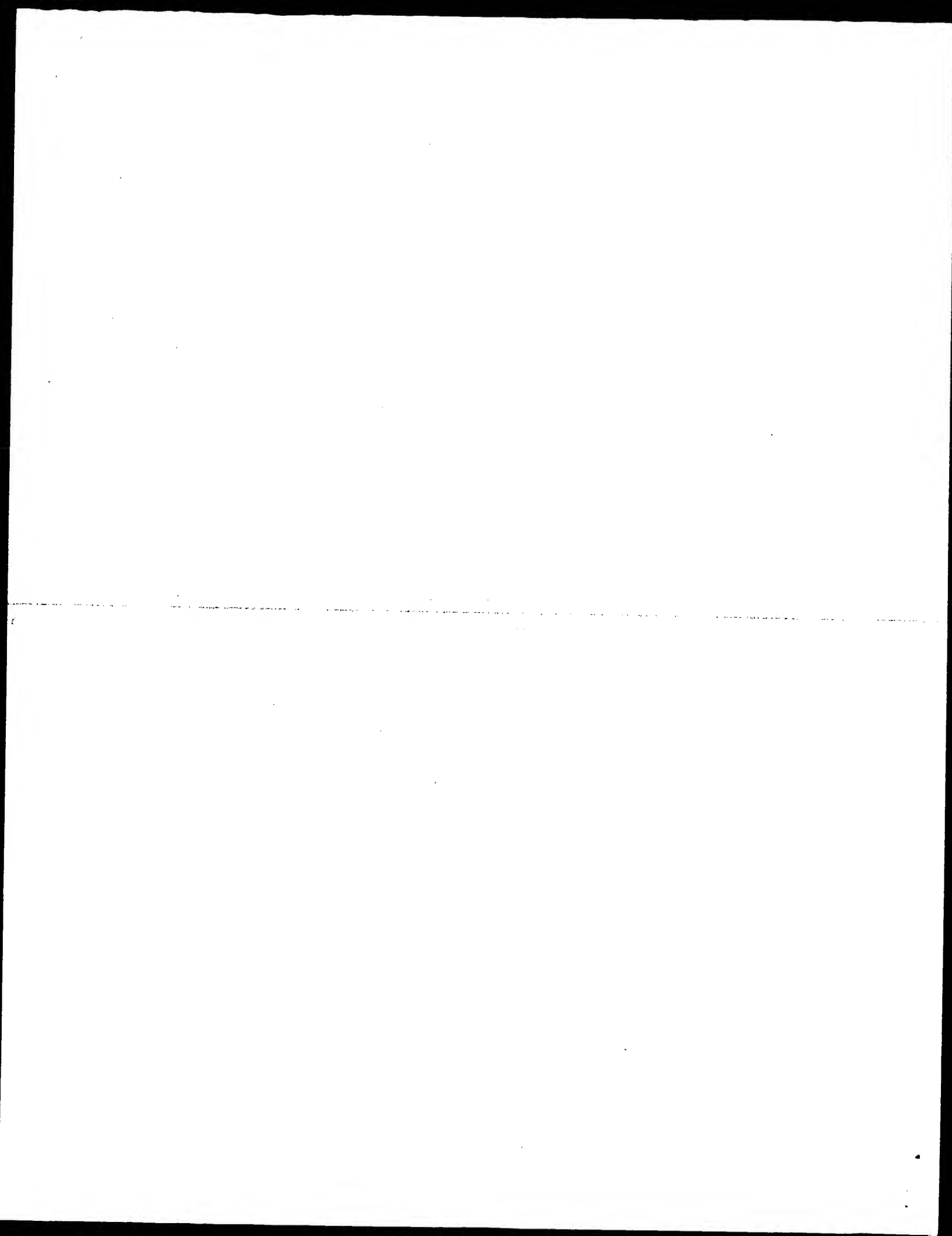
800   TCAACAGCACAAGTGCCTTCGGGAGTCGGGATCTTGGACCGAGTA 849
      ||| |||:||||: :||| ||| :|||:|||||:|||||
288   AlaSpGlyMetLeuPheGlyAlaLeuGlyProCysProValCysAlaAs 304
      |||:|||||:|||||:|||||:|||||:|||||:|||||
850   GCTGATGCGATGGTTCGGGTCCTCTCCCTGCGAGGAATCTCGGG 899
      |||:||||:|||||:|||||:|||||:|||||:|||||
304   nGlyMetTyrTyrTrsAnGlyGlnTyrGlnCysSerGlyAsnValSerG 321
      :||: :||: :||| ||| |||:|||||:|||||:|||||
900   TCAGTGTGTCATTCAAGAGCGATGCTATTACTGACTGGGAGCTCACTG 949
      |||:||||:|||||:|||||:|||||:|||||:|||||
321   LuTrpSerLysCysThrTyrSerAlaThrGluProValAlaArgValLys 337
      |||:|||||:|||||:|||||:|||||:|||||:|||||
950   CTTGACCAAGATGATGTGTAAGACACAGACACCCACCGG.....AAG 993
      |||:|||||:|||||:|||||:|||||:|||||:|||||
338   LysTrpGlnIleProHisGlyThrLysAsnAspTyrLeuMetLysTrpPh 354
      :|||: ||| |||:|||||:|||||:|||||:|||||:|||||
994   GAGTGGGTAAACCCCA.....AAGSATTCCGAGAAATCTCTTACCT 1034
      :|||: ||| |||:|||||:|||||:|||||:|||||:|||||
354   eLysSerGlnLysValLysLysProGluArgValLeuProProMetSer. 370
      :|||: ||| |||:|||||:|||||:|||||:|||||:|||||
1035   CAAGAAATTGAAGTTAAAAAGCAGGACCGCTATATTCCTCCCAAGAACCA 1084
      :|||: ||| |||:|||||:|||||:|||||:|||||:|||||
371   .....ProGluLysSerGlySerLysAla 378
      :|||: ||| |||:|||||:|||||:|||||:|||||:|||||
1085   GCGCCTCGTGGCGGCCACGCGCTCGGCCCTCCACAGCCTCGGCTCTGCT 1134
      :|||: ||| |||:|||||:|||||:|||||:|||||:|||||
379   ThrGlnArgThrSerLeuLeuSerLysGlyLeuAspLysLeuArgPh 395
      :|||: ||| |||:|||||:|||||:|||||:|||||:|||||
1135   GCTGTGAAGTCTCTGCTTCAGCAGATAGCCATTATCCAAACATGAGAT 1184
      :|||: ||| |||:|||||:|||||:|||||:|||||:|||||
395   eSerValValGly.....GlnSerLysGluAlaAlaAsnGluTrpI 409
      :|||: ||| |||:|||||:|||||:|||||:|||||:|||||
1185   CTTGACTCTCGGAAGCTGTCGCGAACAAGGATGAAGTGAAGGCCATGA 1234
      :|||: ||| |||:|||||:|||||:|||||:|||||:|||||
409   LeGluLysLeu.....LysLeuAlaGlyAlaAsnPheTyrAlaArgVal 423
      :|||: ||| |||:|||||:|||||:|||||:|||||:|||||
1235   TTGAAACTCGGGGGAAGTTGACGGGGAGCGGCCAACAGGCTTCCCTG 1284
      :|||: ||| |||:|||||:|||||:|||||:|||||:|||||
424   ValLysAspIleAspCysLeuIleAlaCysGlyGluLeuAspAsnGluAs 440
      :|||: ||| |||:|||||:|||||:|||||:|||||:|||||
1285   .....TGCTCAGCAGCAAAAGAGGAGTGGAAAGATGAA 1319
      :|||: ||| |||:|||||:|||||:|||||:|||||:|||||
440   nAlaGluValArgLysAlaArgArgLeuLysIleProIleValArgGlu 457
      :|||: ||| |||:|||||:|||||:|||||:|||||:|||||
1320   TAAGAGATGGGAAGTAAAGGAAGCCAAATCCGAGTTGTGTCTCAGG 1369
      :|||: ||| |||:|||||:|||||:|||||:|||||:|||||
457   lyTyrIleGlyCysValLysArgThrLys..... 467
      :|||: ||| |||:|||||:|||||:|||||:|||||:|||||
1370   ACTTCTCCAGGACGCTCTCCGCTCCCAAGACGCTTCAGGAGTTGTTTC 1419
      :|||: ||| |||:|||||:|||||:|||||:|||||:|||||
468   CysCysHisLeuIleCysIleAsnTrpAsnAlaLeuGluSerSerLysG 484
      :|||: ||| |||:|||||:|||||:|||||:|||||:|||||
1420   TTAGCGCACATCTTGTC...CTTTGGGGGCGAGAGTGAAGGCGAGACC 1466
      :|||: ||| |||:|||||:|||||:|||||:|||||:|||||
484   y***ThrValThrValLysValLysGlyArgSerAlaCysSer***SerP 501
      :|||: ||| |||:|||||:|||||:|||||:|||||:|||||
1467   TGTTGAAGTT...GTGGCCCCAAGAGGGAAGTCAGGGGCTCGCGCTCTCCA 1513
      :|||: ||| |||:|||||:|||||:|||||:|||||:|||||
501   ro***ValCysLysAsnThrAlaHisIlePro*****Trp..... 514
      :|||: ||| |||:|||||:|||||:|||||:|||||:|||||
1514   AA.....AAAGCAGGGCGCCAGGTCAAGAGGAGGATATCAACAAA 1554
      :|||: ||| |||:|||||:|||||:|||||:|||||:|||||
515   ...GluLysHisIleGlnCys***LeuLys..... 523
      :|||: ||| |||:|||||:|||||:|||||:|||||:|||||
1555   TCTGAAAGAGAGAAATGAAATTAACCTTTAAAGAGAGAGCAGCTGTGGATCC 1604
      :|||: ||| |||:|||||:|||||:|||||:|||||:|||||
524   .....HisValLeu..... 526
      :|||: ||| |||:|||||:|||||:|||||:|||||:|||||
1605   TGATTCTGGACTGGACACTCTGCGCATGTCTCTGAGNAGGTGGGAGG 1654
      :|||: ||| |||:|||||:|||||:|||||:|||||:|||||
527   .....Thr***His***Val.....CysThrGly 534
      :|||: ||| |||:|||||:|||||:|||||:|||||:|||||

```

1655 TCTTCAGTCGCCACCTGGCCCTGGTGACATCGCTTAAGGAACCAACTCC 1704
535 TyrTrpValLeuGluInlelleGluInAspGlySerGluCysTrpVal 551
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1705 TACTACAAGCTGCAGCTTCGTGGAGGACGACAAAGAAACAGGTATTGGAT 1754
551 PheArgLysTrpGlyArgValGlySerGluLysIleGlyGlyGlnLysL 568
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1755 APTCAGGTCTCTGGGCGCTGTGGGTAG...GTGATCGGTAGCAACAAC 1801
568 euGluGluMet...SerLysThrGluAlaIleLysGluPheLysArgLeu 583
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1802 TGGAAACAGTCCCTCAAGAGAGATGCCATTGAGCAGTTTCATGAATTA 1851
584 PheLeuGluLysThrGlyAsnSerTrpGluAlaTrpGluCysLysThrAs 600
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1852 TATGAAGAAACCGGACGCTTGGCACTCC.....AAAA 1889
600 nPheArgLysGlnProGlyArgPheTrpProLeuAspValAspTrpGlyV 617
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1890 TTTCACGAAGTATCCCAAAAGTTTACCCTCTGGAGATTGACTATGGCC 1939
617 al...LysLysAlaProLysArgLysAspIleSerGlu...MetLysSer 631
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1940 AGGATGAAGAGCGAGTGAAGAGGTCACAGTAAATCCTGGCACCAAGTCC 1989
632 SerLeuAlaProGlnLeuGluLeuMetLysMetLeuPheAsnValG1 648
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1990 AAGCTCCCAAGCCAGTTCAGGACCTCATCAGATGATCTTTGATGTGA 2039
648 uThrTrpArgAlaAlaMetMetGluPheGlu***AsnMetSerGluMetP 665
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2040 AAGTATGAAGAAAGCCATGGTGGATGATGAGATCGACCTTCAGAAATGC 2089
665 roLeuLysLysLeuSer*****AsnIleGlu***GlyPheGluAlaLeu 681
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2090 CTTGGGGAAGCTGAGCAAGAGCAGATCCAGCGCATCTACTCCATCCTC 2139
682 Thr*****LeuPheGluGlyHisArg***SerSerThrGlyLe 698
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2140 AGT.....GAGTCCACAGCGCGTCTCTCAGGGCAG 2171
698 u***GluLysAla***LeuLeu*****PheSerLeuLeu 715
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2172 CAGCGACTCTCAGATCCTCGATCTCAAAATCGCTTTTACACCTGATCC 2221
715 erLeuLeuPheIleLeuLeuTrpGlyMetArgMet..... 727
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2222 CCCACGACTTT.....GGATGAAGAGCCTCCGCTCCTG 2256
728IleSerTyrSerLysAlaLysMetLeuGluAlaLeuG1 740
2257 AACATCGACAGATCTGACGGCCAGGTGGAAATGCTTGACAACTGCT 2306
740 nAspIleGluAlaSerLysIleVal.....GlyPheAspSerAsps 755
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2307 GGACATCGAGTGGCTTACAGTCTGCTCAGGGAGGGTCTGATGATGCA 2356
755 erAspGluSerLeuAspAspLysTrpMetLysLeuHisCysAspIleThr 771
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2357 GCAAGGATCCCATGATGTCAACTATGAGAAGCTCAAAACTCACATTAA 2406
772 ProLeuAlaHisAspSerGluAspTrpLysLeuIleGluGlnTrpLe 788
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2407 GTGGTTGACAGAGATCTGAAGAGCCGAGATCATCAGGAGATGTGTAA 2456
788 uAsnThrHisAlaProThrHisLysAspTrpSerLeuLeuLeuGluGluV 805
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2457 GAACACTCATGCCAACACACACAGTCCGTATGACTTGGAGTCATCGATA 2506
805 alPheSerLeuAspArgAspGlyGluLeuAsnLysTrpSerArgTrpLys 821
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
507 TCTTTAAGATGAGCGTGAAGGCAATGCCAGCGTTTCAAGCCCTTTAAG 2556

us-09-236-995d-2.rge

Mon Mar 11 09:58:37 2002




```

1 AsnlysmetleuLeuTrpHisGlySerArgLeuThrAspPheValGlyI11 17
|||||
2549 AACAGAGAGCTATTATGCGACGCTTCAAGGTGACGAATTTTGGGGAAT 2598
17 eleuSerGlnGlyLeuArgIleAlaProGluAlaProValThrGlyT 34
|||||
2599 TCTTAGTCAAGGGCTAAGATTGCACTCCCGAGGACCTCTTCTGCGCT 2648
34 yrmelPheGlyLysGlyLeuTyrrPheAlaAspLeuValSerLysSerAla 50
|||||
2649 ATATGTCGCGAAAGGCTCTACTTTCGAGATCTAGTACCAAGAGCGCA 2698
51 GlnTyrrCysTyrrValAspArgAsnAsnProValGlyLeuMetLeuLeu 67
|||||
2699 CAATACCTGTTATGTGATAGGAATATCTGTAGCTTTGATGCTTCTTTC 2748
67 rgluValAlaLeuGlyAspMetTyrrGluLeuLysLysAlaThrSerMet 84
|||||
2749 TGAGGTCCTTAGAGAGCATGTATGAACCTAAGAACCCGTCATG 2798
84 sPLysProProArgGlyLysHisSerThrLysGlyLeuGlyLysThrVal 100
|||||
2799 ACAAACTCCCAAGAGGAGCATTCGACCAAGGATTAGCGCAAAACCGTG 2848
101 ProLeuGlnSerGluPheValLysTrpArgAspValValProGly 117
|||||
2849 CCACGTGAGTCAGAGTTGTGAGTGGAGGATGATGTGAGTTCCTG 2898
117 sGlyLysProValProSerSerIleArgSerSerGluLeuMetTyrrAsn 134
|||||
2899 CGGCAAGCGGTCATCATCATATAGAGCTCGAAGCTCATGTACATG 2948
134 luTyrrIleValTyrrAsnThrSerGlnValLysMetGlnPheLeuLys 150
|||||
2949 AGTACATGCTCTACACACATCCCGAGTGAAGATGCACTTCTGCTGAG 2998
151 ValArgPheHisHisLysArg 157
|||||
2999 GTGCTTTCATCACAAGAGG 3019
seq_name: gb_P1:AF093627
seq_documentation_block:
LOCUS AF093627 3285 bp mRNA PLN 29-NOV-1998
DEFINITION Zea mays poly(ADP)-ribose polymerase (PARP1) mRNA, complete cds.
ACCESSION AF093627
VERSION AF093627.1 GI:3928870
KEYWORDS
SOURCE
ORGANISM
Zea mays.
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 3285)
REFERENCE
AUTHORS Mahajan,P.B. and Zhu,Z.
TITLE Purification and cDNA cloning of maize Poly(ADP)-ribose polymerase
JOURNAL Plant Physiol. 118 (3), 895-905 (1998)
MEDLINE 99026291
REFERENCE
AUTHORS Mahajan,P.B. and Zhu,Z.
TITLE Direct Submision
JOURNAL Submitted (22-SEP-1998) Crop Protection, Pioneer HiBred
International, Inc., 7300 NW 62nd Avenue, Johnston, IA 50131-1004,
USA
FEATURES
SOURCE
1. .3285 Location/Qualifiers
/organism="Zea mays"
/db_xref="taxon:4577"
1. .3285
/gene="PARP1"
100. .3042
/gene="PARP1"
/EC_number="2.4.2.30"

```

```

/function="add ADP-ribose"
/codon_start=1
/product="poly(ADP)-ribose polymerase"
/protein_id="AAC79704.1"
/db_xref="GI:3928871"
/translation="MAAPPKMAKVEYAKSGRASCSRSPIARDQRLGKMKVASQFD
GFEMPMNARCIFSKNOKISVDVEGIDALRMDDEKIRNRYGSAAGSTPAAPPE
KCTEIAIPASRCSRCSEKITKGSVLEAKLESEGRGIPMYHANCFFVSPAAVE
KFSGMPTLDEDRKTMILYKDDVGNKQKSKSENDISYSARLDESGDT
VRNGQLVDPKGSSTSSADTOLKIRESDPLMKLKKLTHVSAALRDLKLEANGDT
SGPRHLILDRADDMFLGALGPCVCAKMGMTYINGQIQCGANVSMSKCTYSATPEYR
VKRWQIPHGTRKNDYLMKWFKSKQVKKPERVLPMSPEKSGSATORTSLSSGLDK
LRFVSGQSPKKAANEMIEKLKLAGANFYARVVDIDCLICGELDENAVRARRLK
IPIVREGIGECVKKNMKLPDILYKLENALESSKSTVYKVRGSRVHSSGLDTA
HILDECKSTVMAATIMSDIALGVNSYVLIQIIBDDGSECEYFRKMGVSEKIGQK
LEBKSITBAIKERFRLPLEKTGNSEMECKTNRQGRFPLVDUDYGVKAPKRRD
ISEKSSILAPLELMKLEFVETRYRAMMEFETINMSMPGLKSKENIRKGEALTE
IONLKDTPADALAVRSLIVAAASNREFTLIPSTIHPHIRDEDLMTKAMBALDOI
EIASKIVGFDSDSDSLDDKMKLHCDITPLADSEDPYKLIEDYLTNTHAPTRKMSL
ELEEVFSLDRGELNKYSRYRKNLHNMKMLHMGSRLEPFGIISQGLRIAPPAVYTG
YMRKGLYFADIVSKSAOYCVVDNRNPNVGLMLSEVALGDMVELKATKATSMDKPRKH
STKGLGKTVPLESEFVKRRDVTVPCKPVPSSIRSELWYNEIYVNTSQVAKMQLL
KVRPHHKR"

```

```

alignment_scores:
Quality: 828.00 Length: 157
Ratio: 5.274 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

```

alignment_block:

US-09-236-995d-4 x AF093627 ..

Align seq 1/1 to: AF093627 from: 1 to: 3285

```

1 AsnlysmetleuLeuTrpHisGlySerArgLeuThrAspPheValGlyI11 17
|||||
2569 AACAGATGCTATTATGCGACGCTTCAAGGTGACGAATTTTGGGGAAT 2618
17 eleuSerGlnGlyLeuArgIleAlaProGluAlaProValThrGlyT 34
|||||
2619 TCTTAGTCAAGGGCTAAGATTGCACTCCCGAGGACCTCTTCTGCGCT 2668
34 yrmelPheGlyLysGlyLeuTyrrPheAlaAspLeuValSerLysSerAla 50
|||||
2669 ATATGTCGCGAAAGGCTCTACTTTCGAGATCTAGTACCAAGAGCGCA 2718
51 GlnTyrrCysTyrrValAspArgAsnAsnProValGlyLeuMetLeuLeu 67
|||||
2719 CAATACCTGTTATGTGATAGGAATATCTGTAGCTTTGATGCTTCTTTC 2768
67 rgluValAlaLeuGlyAspMetTyrrGluLeuLysLysAlaThrSerMet 84
|||||
2769 TGAGGTCCTTAGAGAGCATGTATGAACCTAAGAACCCGTCATG 2818
84 sPLysProProArgGlyLysHisSerThrLysGlyLeuGlyLysThrVal 100
|||||
2819 ACAAACTCCCAAGAGGAGCATTCGACCAAGGATTAGCGCAAAACCGTG 2868
101 ProLeuGlnSerGluPheValLysTrpArgAspValValProGly 117
|||||
2869 CCACGTGAGTCAGAGTTGTGAGTGGAGGATGATGTGAGTTCCTG 2918
117 sGlyLysProValProSerSerIleArgSerSerGluLeuMetTyrrAsn 134
|||||
2919 CGGCAAGCGGTCATCATCATATAGAGCTCGAAGCTCATGTACATG 2968
134 luTyrrIleValTyrrAsnThrSerGlnValLysMetGlnPheLeuLys 150
|||||
2969 AGTACATGCTCTACACACATCCCGAGTGAAGATGCACTTCTGCTGAG 3018

```

151 ValArgPheHisLysArg 157
 |||||
 3019 GTGGTTTCCATCAAGAGG 3039

seq_name: gb_pl:ATH131705

seq_documentation_block:

LOCUS ATH131705 3187 bp mRNA PLN 18-DEC-1998
 DEFINITION Arabidopsis thaliana mRNA for poly(ADP-ribose) polymerase.

ACCESSION AJ131705

VERSION AJ131705.1 GI:4038490

KEYWORDS NAD(+) ADP-ribosyltransferase; parp-1 gene; poly(ADP-ribose) polymerase.

SOURCE thale cress.

ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 1 (bases 1 to 3187)

REFERENCE

AUTHORS kazmaier, M.

TITLE Direct Submission

JOURNAL Submitted (17-DEC-1998) Kazmaier M., Departement D'Ecophysiologie Vegetale Et De Microbiologie, Commissariat A L'Energie Atomique, CEA Cadarache, Lab. De Radiobiol. Vegetale, Bat. 185, 13108 St. Paul-les-Durance, FRANCE

REFERENCE

AUTHORS Doucet-chabeaud, G. and Kazmaier, M.

JOURNAL Unpublished

FEATURES Location/Qualifiers

source

1..3187

/organism="Arabidopsis thaliana"

/cultivar="landsberg erecta"

/db_xref="taxon:3702"

11..2963

/gene="parp-1"

11..2963

/gene="parp-1"

/standard_name="NAD(+) ADP-ribosyltransferase"

/EC_number="2.4.2.30"

/function="ADP-Ribose polymer synthesis"

/note="mol. weight: 111 kDa"

/codon_start=1

/evidence="experimental"

/product="poly(ADP-ribose) polymerase"

/protein_id="CAA10482.1"

/db_xref="GI:4038491"

/translation="MASPHKPPWRAEYAKSSRSCKTCKSVINKENFRLGKLVOSTHFD
 GIMPMWNAHSCILAKTKQIKSVDDVEGIESLRWEDQOKIRKYVESGAGSNTSTGHS
 TSSTANNAKLEYGLEVSQTSRAGCRKCEKILGVEIRIFSKPEGPNGKGLMWHAKCF
 LEMSSTELSLGWSRIPDSQDQALLPLVKKALPAAKETAARQTNSTRAGTKRND
 SYDNEKSLAKSSFDMSGALQPCSEKEEQAQTKELWDLKDDLKRYVTSAELREML
 EVNQSTRGSLDLDRKADGMFGLALCPMGSHLSFGGLYRCHCYISEWSKSH
 STLDPRIKGKKWIPDETENOFLLKWNKSOKSVKPKRILRPVLSGETSOGGSKDAD
 SSRERLADLKVSTAGNKKRPQPKKPIEAGAEFHANVKKGTSCLVVCGSLDIRDAE
 MKRARKVAIRVDYLVDFCKQKPLPFKIKETIDETSESLVTIVKGRSAVHEASGL
 QERCHILGDNSTLTLSMSDSTGINSYIILOIQEDKGSDCYFVRKGRVGNKFI
 GKNVKEESKSDAVHEFKRLFLKGTNTWSEQKTNFQPGKFLPIDIDYGVNKFV
 AKKPFQTSNAPSLIELMLKMLFDVETYSAMEFETINSEMPGLKSLKHNOKQFE
 ALTIQRLTSEDPQPTMESLVDASNRFETIPSTPHIIRDEDDFKSKVMLEAL
 QDIASRIKGFVDVSTESLDKYLKLCDDISPLPHDSERYRLIEKVLNTHAPTHE
 WSLREVFVLEEREGEDKYPAPHEKLCNKLMLHWSRNLNFVGLNQGLRIAPPEAP
 ATGVMFGKGIYFADLVSKAQCYCTCKNPVLMJLSEVALGEIHELTRAKYMDKRPFR
 GKHSKGLGKRPQDSEFAKWRGDTVPCGKPVSSKVKASELMYNEYIVYDTAQVQLQ
 FLKVRFEKRR"

mat_peptide

14..2959

/gene="parp-1"

/EC_number="2.4.2.30"

/product="poly(ADP-ribose) polymerase"

1044 a 577 c 743 g 823 t

BASE COUNT

ORIGIN

alignment_scores:

Quality: 697.00 Length: 157
 Ratio: 4.709 Gaps: 0
 Percent Similarity: 94.268 Percent Identity: 82.166

alignment_block:

US-09-236-995D-4 x ATH131705 ..

Align seg 1/1 to: ATH131705 from: 1 to: 3187

1 AsnLysMetLeuLeuThrPhisGlySerArgLeuThrAsnPheValGlyI 17
 |||||
 2489 AATAGATGCTCTCTATGGCATGGTTCTCGATTAAAGCAATTTTGTGGAAT 2538

17 eLeuSerGlnGlyLeuArgIleAlaProProGluAlaProValThrGlyT 34
 |||||
 2539 ATTGAACCAAGACATGAGAAATTCACCTCCAGAACTCTGCTACTGGTT 2588

34 yMetPheGlyLysGlyLeuThrPheAlaAspLeuValSerLysSerAla 50
 |||||
 2589 ACATGTTTGGAAAGGATATATCTTTGCTGACCTTGTCTAGTAAAGTGTCT 2638

51 GlnTyrCysTyrValAspArgAsnAsnProValGlyLeuMetLeuLeu 67
 |||||
 2639 CAGTACTGCTACACTGTGAAGAAAAATCCGGTGGTCTAATGCTCTCTGAG 2688

67 rGluValAlaLeuGlyAspMetTyrGluLeuLysLysAlaThrSerMetA 84
 |||||
 2689 TGAAGTTGATGGGAGAAATACATGAGCTACAAAAGCTAAGTATATGG 2738

84 sPLysProArgGlyLysHisSerThrLysGlyLeuGlyLysThrVal 100
 |||||
 2739 ATAAACCTCCGAGAGGAAACACATCGACCAAGGGCTCGCAAGAAAGTG 2788

101 ProLeuGluSerGluPheValLysThrPargAspValValProCy 117
 |||||
 2789 CCTCAAGATTCCGAGTTTCCCAAGTGGAGAGTGATGTGACTGTTCCTCG 2838

117 sGlyLysProValProSerSerIleArgSerSerGluLeuMetTyrAsnG 134
 |||||
 2839 TGGAAACCTGTTTCATCAAGGTCAAGGCTTCGAGCTTATGTACAAATG 2888

134 luTyrIleValTyrAsnThrSerGlnValLysMetGlnPheLeuLys 150
 |||||
 2889 AGTATATGCTCTACGATACAGCCAGGTGAAGTTGCAGTTCTTGTGAAA 2938

151 ValArgPheHisHisLysArg 157
 |||||
 2939 GTAAGTTTAAAGCACACAGAGA 2959

seq_name: gb_ro:RNPARP2

seq_documentation_block:

LOCUS RNPARP2 1939 bp DNA ROD 20-MAY-1992

DEFINITION R.norvegicus mRNA for poly(ADP-ribose) polymerase (clone PRATC).

ACCESSION X65497

VERSION X65497.1 GI:56849

KEYWORDS NAD(+) ADP-ribosyltransferase.

SOURCE Norway rat.

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

REFERENCE 1 (bases 1 to 1939)

AUTHORS Thibodeau, J., Gradwohl, G., Dumas, C., Clairoux-Moreau, S., Brunet, G.,

Penning, C., Poirier, G.G. and Moreau, P.

TITLE Cloning of rodent cDNA coding the poly(ADP-ribose) polymerase

JOURNAL catalytic domain and analysis of mRNA levels during the cell cycle

MEDLINE Biochem. Cell Biol. 67 (9), 653-660 (1989)

REFERENCE 90027702

AUTHORS 2 (bases 1 to 1939)

Potvin, F.

TITLE Direct Submission

JOURNAL Submitted (23-MAR-1992) F. Potvin, Molecular Endocrinology, CHUL

Research Center, 2705, Boul Laurier, Ste-Foy, Quebec, G1V 4G2,
CANADA
See also X55496.
Location/Qualifiers
1. 1939
/organism="Rattus norvegicus"
/strain="Sprague Dawley"
/db_xref="taxon:10116"
/chromosome="1q41-42"
/tissue_type="prostate"
/clone_lib="lambda EMBL4"
/clone="PRATC"
32. 1510
/EC_number="2.4.2.30"
/codon_start=1
/product="Poly(ADP-ribose) polymerase"
/protein_id="CAA46478.1"
/db_xref="GI:56850"
/db_xref="SWISS-PROT:P27008"
/translation="MKLTGLGGAADPDGSLHSAHVLEKGGKVFSAITGLVDIVKGT
NSYKIQLESDESKSRIWFRSMRGVTVIGSNKLEQMSKEDDAVEHFMKLYEETGN
AWHSKNEYTPKPFHPLAIDYQDEEAVKLAQVPGTKSLKLPVQVLEGMFLDVEISM
KKALVEYEDIQMLPKLSRRQIAQYSLSEVOQAVSQGSESDILDSNRFYTLI
PHDFGMMKPDLENTDSQVAKVEMLDNLLDIEVAYSLRGGSDSDSKDPIDVNYEKLK
TDIKVVDROSEAEVIRKYNTHATTHAYDLVIDIFKIEREGESQRYKPFROLHN
RLLWHSRTNFGILSGLRIAPPEAPVGYMGKGIYFADMVSKSANYCHTSQGD
PIGLILGEVALGNMVELKHASHISKLPGKHSVGLGKTADPSPASITLDGVEVPLG
TGIPSGVNDTCLLYNEIYVDIAQVNLKLLKLFNFKTSLW"

BASE COUNT 524 a 468 c 521 g 426 t
ORIGIN

alignment_scores:
Quality: 522.50 Length: 156
Ratio: 3.842 Gaps: 1
Percent Similarity: 87.179 Percent Identity: 62.821

alignment_block:
US-09-236-995d-4 x RNPAP2

Align seg 1/1 to: RNPAP2 from: 1 to: 1939

1 AsnLysMetLeuLeuThrHisGlySerArgLeuThrAsnPheValGlyI1 17
1031 AACCGAGACTCTCTGGCAGCGGTCCAGACACCACCTTCGAGCAT 1080
17 eLeuSerGlnGlyLeuArgGlyAlaProProGluAlaProValThrGlyT 34
1081 CCGTGCACAGGCTCTCGGATAGCCGCCACCTGAAGCACCCTGTGACAGGCT 1130
34 yMetPheGlyLysGlyLeuTyrPheAlaAspLeuValSerLysSerAla 50
1131 ACATGTTTGGGAAGAAATCTACTTTCGTGATATGCTGTCACAAAGTGG 1180
51 GlnTyrCysTyrValAspArgAsnAnProValGlyLeuMetLeuLeuSe 67
1181 AACTACTGCCACACGCTCAGGAGACCCGATGCTTAATCTGTTGG 1230
67 rGluValAlaLeuGlyAspMetTyrGluLeuLysLysAlaThrSerMetA 84
1231 AGAAGTTGCCCTTGGAAACATGTACCAACTCAAGCATGCTTCACATCA 1280
84 sPLysProProArgGlyLysHisSerThrLysGlyLeuGlyLysThrVal 100
1281 GCAAGTTACCAAGGCAACGACAGTGTCAAGGTTTGGGCAAAACCGCC 1330
101 ProLeuGluSerGluPheValLysThrArgAspValValValProCy 117
1331 CCTGACCTTCGCCAGCATCAC...CTGATGGTGTAGAGGTTCCGCT 1377
117 sGlyLysProValProSerSerIleArgSerSerGluLeuMetTyrAsnG 134
1378 GGAACAGGAGATCCCTCTGGTGTAAATGACACACCTCCCTGCTGCTATAACG 1427

134 LuTyrIleValTyrAsnThrSerGlnValLysMetGlnPheLeuLeuLys 150
1428 AGTACATTTGCTACGACATGCTCAGGTGATCTGAAGTACCTACTGAAA 1477
151 ValArgPheHisLys 156
1478 CTGAAGTTCAATTTCAAG 1495
seq_name: gb_ro:RNU94340
seq_documentation_block:
LOCUS RNU94340 3128 bp mRNA ROD 19-FEB-1998
DEFINITION Rattus norvegicus poly(ADP-ribose) polymerase mRNA, complete cds.
ACCESSION U94340
VERSION U94340.1 GI:2896791
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 3128)
Benke.S., Meyer.R. and Burkle.A.
Isolation of cDNA encoding full-length rat (Rattus norvegicus) poly
(ADP-ribose) polymerase
Biochem. Mol. Biol. Int. 43 (4), 755-761 (1997)
89046546
2 (bases 1 to 3128)
Benke.S., Meyer.R. and Burkle.A.
Direct Submission
Submitted (18-MAR-1997) Angewandte Tumorstudiologie, Deutsches
Krebsforschungszentrum, Im Neuenheimer Feld 242, Heidelberg
D-69120, Germany
3 (bases 1 to 3128)
Benke.S., Meyer.R. and Burkle.A.
Direct Submission
Submitted (19-FEB-1998) Angewandte Tumorstudiologie, Deutsches
Krebsforschungszentrum, Im Neuenheimer Feld 242, Heidelberg
D-69120, Germany
Sequence update by submitter
On Feb 19, 1998 this sequence version replaced gi:2583143.
Location/Qualifiers
1. 3128
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/cell_type="mononuclear blood cell"
31..3075
/codon_start=1
/product="poly(ADP-ribose) polymerase"
/protein_id="AAC53544.1"
/db_xref="GI:2896792"
/translation="NAEATERYRYEYAKSGRACKCKSESIPKDSLRMTMVQSPMF
DGKVFPHWTFSCFWKVGHSIRPDTEVDGFESELDWDOOKVKKTAEAGVAGKGGHGG
GGAETLGDFAAETAKSNRSTCKGCMERKEQKQMLSKMLDPEKPOLGMDRWHYP
TCVKNRDELGRPEYSSQLGFSLLSAEDKEALKQLPAVKSSEGRKRCDEVDIGE
VAKKSKGKDESSKLEKALQANLWNTIDELKACSTNDLKEILLIFNOQVPSG
ESAILDRVADGMAGALLPKCKCQGLVFKSDAYCTGDTAVTKMVKTNQPSRKE
VTPKFEISYLKLIKIKKDRLPFPSSAPAPPAPPVISITSAPTAVNSKAPADKPLS
NMKILTLKLSNKKDEAKMIEKLGKLTGSANKASLCISTKKEVKMSKMEEVKAA
NVRVCEDFLDQVDSASLSQLGALLSAHSLSSWGAEVKVPEGVVVKCKSAAPSKSK
GAVKCEVGNKSEKRMKLTLLGGAADPDGSLHSAHVLEKGGKVFSAITGLVDIVKGT
NSYKIQLESDESKSRIWFRSMRGVTVIGSNKLEQMSKEDDAVEHFMKLYEETGN
AWHSKNEYTPKPFHPLAIDYQDEEAVKLAQVPGTKSLKLPVQVLEGMFLDVEISM
KKALVEYEDIQMLPKLSRRQIAQYSLSEVOQAVSQGSESDILDSNRFYTLI
PHDFGMMKPDLENTDSQVAKVEMLDNLLDIEVAYSLRGGSDSDSKDPIDVNYEKLK
TDIKVVDROSEAEVIRKYNTHATTHAYDLVIDIFKIEREGESQRYKPFROLHN
RLLWHSRTNFGILSGLRIAPPEAPVGYMGKGIYFADMVSKSANYCHTSQGD
PIGLILGEVALGNMVELKHASHISKLPGKHSVGLGKTADPSPASITLDGVEVPLG
TGIPSGVNDTCLLYNEIYVDIAQVNLKLLKLFNFKTSLW"

BASE COUNT 861 a 758 c 905 g 604 t
ORIGIN

```

alignment_scores:
  Quality: 522.50      Length: 156
  Ratio: 3.842        Gaps: 1
  Percent Similarity: 87.179  Percent Identity: 62.821

alignment_block:
  US-09-236-995D-4 x RNU94340

  Align seg 1/1 to: RNU94340 from: 1 to: 3128

1 AsnLysMetLeuTyrHisGlySerArgLeuThrAsnValGlyI1 17
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
2596 AACCGGAGACTGCTGCGCAGGGTCCAGGACCCACCACTTCGACAGCAT 2645

17 eLeuSerGlnGlyLeuArgIleAlaProGluAlaProValThrGlyT 34
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
2646 CCTGTCACAGGGTCTCGGATAGCCACCTGAAGCACCCTGTCACAGGCT 2695

34 yrMetPheGlyLysGlyLeuTyrPheAlaAspLeuValSerLysSerAla 50
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
2696 ACATGTTGGGAAGGAATCTACTTTGCTGATATGTTGCCAAAGTGCG 2745

51 GlnTyrCysTyrValAspArgAsnAsnProValGlyLeuMetLeuLeuSe 67
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
2746 AACTAGTCCACAGCTCTCAGGAGACCCGATGCTTAATCTGTTGGG 2795

67 rGluValAlaLeuGlyAspMetTyrGluLeuLysLysAlaThrSerMeta 84
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
2796 AGAAGTTGCCCTTGGAAACATGTACGAACCTCAAGCATGCTTCTCATCA 2845

84 sPlyProProArgGlyLysHisSerThrLysGlyLeuGlyThrVal 100
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
2846 GCAAGTTACCCAGGCGCAGCAGTGTCAAGAGTTTGGGCAAAACCGCG 2895

101 ProLeuGluSerGluPheValLysTyrArgAspAspValValProCy 117
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
2896 CCTCACCTCGGCCAGCATCAC...CTGGATGGTGTAGAGTTCCTGCT 2942

117 sGlyLysProValProSerSerIleArgSerGluLeuMetTyrAsnG 134
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
2943 GGAACAGGAGTTCCTGCTGTTAATGACACCTGCCTGCTGATAACG 2992

134 luTyrIleValTyrAsnThrSerGlnValLysMetGlnPheLeuLeuLys 150
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
2993 AGTACATTGCTACGACATGCTCAGTGAATCTGAAGTACCTACTGAAA 3042

151 ValArgPheHisHisLys 156
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
3043 CTGAAGTCAATTTCAG 3060

seq_name: gb_ro:MMADPRP

seq_documentation_block:
LOCUS MMADPRP 3172 bp mRNA ROD 12-SEP-1993
DEFINITION Mouse mRNA for poly (ADP-ribose) polymerase (BC 2.4.2.30).
ACCESSION X14206
VERSION X14206.1 GI:49893
KEYWORDS DNA-binding protein; NAD(+) ADP-ribosyltransferase; zinc finger protein; house mouse.
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS Huppi, K.
TITLE Direct Submission
JOURNAL Submitted (26-JAN-1989) Huppi K., NCI/NIH, Lab of Genetics, Building 37, Room 2B-21, Bethesda, MD 20892
REFERENCE 2 (bases 1 to 3172)
AUTHORS Evers, R., Hammer, A., and Cornelissen, A.W.C.A.
TITLE Unusual C-terminal domain of the largest subunit of RNA polymerase

```

```

II of Chithidia fasciculata
Nucleic Acids Res. 17, 3387-3401 (1989)
Location/Qualifiers
  1. 3172
  /organism="Mus musculus"
  /strain="BXSB"
  /db_xref="taxon:10090"
  /tissue_type="spleen"
  /clone="pADP3-2.1 and pADP-3C"
  85..3126
  /note="poly (ADP-ribose) polymerase (AA 1 - 1013)"
  /codon_start=1
  /protein_id="CAA32421.1"
  /db_xref="GI:49894"
  /db_xref="MGD:87933"
  /db_xref="SWISS-PROT:P11103"
  /translation="MAFASRLYRVQYAKSGRASCCKSCSIPKDSERMAIMVQSPMF
  DGKVPWHYHSCFWKQGSIRHPDVEYDFSELRWDDQKRTKTBAGGAGVAGQDGS
  GGKAEKTLGDFAAEYAKSNRSMCKGLEIKEGOMRLSKMVDPERQIIGMIDRWYHP
  TCFVKRDELGFREYSASQLGSLISADKRLKQLPAIKNEGRKGDVDTGTHDE
  VAKKSRKTDYKSKLEKALKQNELIWNIDELKKACSTNDLKEILLFNOQOVPSGE
  SAILDRVADGMAGFALLPCKECSQLVFKSDAYCTGDTATMTKCMVKTONPSRKEWV
  TPKEFEILSYLKKLVKKQDRIPPPSSAPITVHWPLSVTSAPTAVNSAPADKPLSN
  MKILTLGKLSQNKDEKAVLEKLGKLTGSKANKASICISIKKEVKMKMKMEVKEAN
  IRVSEDFLDQVSASTKSQDLSAHSLSWPAEVAEPEVVPAPRGSAAPSKSKG
  CFKEGVNKSERKMLTLGGAAVDPDGLSEHSAHLEKGVKFSFATLGLVDIVKGTN
  SYKQLLEDDEKESRYWIFRSWGLGTIGSNKLEQMPSEKEAVEQFMKLEKTKNA
  WHSKNTDYPKFPLEIDYGOEEAVKLTVPKPTGSKLPKPVOLVGMIFDVDSMK
  KALVEIKDLQKMPGLKLSRROQAAYSILSEVQVPSGSSQLDLSNRFTYLIIP
  HDFMKRPLLNADSVQAKVEMLDNLDIEVAISLLRGSDSSKDPIDVRYEKLTP
  DIKVDRESEAEVIRKYKNTHTTHAYDLEVIDIFKIEREGESQRYKPRQLNHR
  RLLHWSRTTFAGILSQGLRIAPPEAPVTGYMFGKGYFADVMVSKSNVCHTSQDDP
  IGLIMLGEVALGNMVELKHAHSIKLPKKGKSVKGLGKTTDPDSASITILEGVEVPLGT
  GIPSGVNDTALLYNEIVYDIAQVNLKLLKLFNFKTSLW"
  misc_feature 88..678
  /note="zinc finger domains"
  misc_feature 2743..2823
  /note="DNA-binding domain A"
  misc_feature 2898..2952
  /note="DNA-binding domain B"
  BASE COUNT 869 a 779 c 925 g 599 t
  ORIGIN

alignment_scores:
  Quality: 519.50      Length: 156
  Ratio: 3.820        Gaps: 1
  Percent Similarity: 87.179  Percent Identity: 61.538

alignment_block:
  US-09-236-995D-4 x MMADPRP

  Align seg 1/1 to: MMADPRP from: 1 to: 3172

1 AsnLysMetLeuTyrHisGlySerArgLeuThrAsnValGlyI1 17
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
2647 AACCGGAGCTGCTGCGCAGGGTCCAGGACCCACCACTTGTCTGTCGAT 2696

17 eLeuSerGlnGlyLeuArgIleAlaProGluAlaProValThrGlyT 34
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
2697 CCTGTCGAGGGTCTGCGGATAGCCACCTGAAGCGCCGCTGACAGGCT 2746

34 yrMetPheGlyLysGlyLeuTyrPheAlaAspLeuValSerLysSerAla 50
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
2747 ACATGTTGGGAAGGATCTACTTTGCCGACATGCTGCCAAAGTGCA 2796

51 GlnTyrCysTyrValAspArgAsnAsnProValGlyLeuMetLeuLeuSe 67
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
2797 AACTACTGCCACATCTCAGGAGACCCGATGCTTAATGCTGTTGGG 2846

67 rGluValAlaLeuGlyAspMetTyrGluLeuLysLysAlaThrSerMeta 84
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
2847 AGAGTTGCCCTTGGAAATATGTATGAACCAAGCATGCTTCACATCA 2896

```

1

REFERENCE
AUTHORS
TITLE
JOURNAL

1 (bases 1 to 3845)
Strausberg, R.
Direct Submission
Submitted (30-JUL-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK
COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: villalon@bcm.tmc.edu
Villalon, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia,
A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,
Muzny, D.M., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 5 Row: p Column: 3
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 49893.

FEATURES
source

1..3845
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="MGC:6498 IMAGE:2648390"
/tissue_type="Mammary tumor. Metallothionien-TGF alpha
model. 10 month old virgin mouse. Taken by biopsy."
/clone_lib="NCI-CGAP_Mam1"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
60..3104
/codon_start=1
/product="Similar to ADP-ribosyltransferase (NAD+; poly
(ADP-ribose) polymerase)"
/protein_id="AAH12041.1"
/db_xref="GI:15080598"
/translation="MAEASERLYRVEYAKSGRACKKSESIPKDSLRMAIMVQSPMF
DKGVHWYHFCFWKGVHSIROPDVEVDGSELWDDQKVKKTAEGVAGKGDGDS
GKAETLGDFAEYAKSNRSMCKGCKLEKQMLSKKWKVDPEKPOLGMDIRWHP
TCFVKRDELGFPEYSQSLKGFSLSAEDREALKQLPAIKNEGKRGKDEVDGTE
VAKKSKGKDKSSKLEKAKNELIWNIDELKACSTNDLKELLIFNQOVPFG
ESAILDRVADGMAFGALLPCKECSQGVFKSDYYCTGDVTATKMTKQNPSPKEW
VTPKEFRTSYLLKLVKKODRIFPPSSAPAPLALPLSVTSAPTAVNSAPADRPKLS
NKKILTGLKSLONKDEKAVIEKLGKLTGSANKASLCISTKKEVKSKMEEVKAA
NVRVCEDFLDVSASTKSLQELLSAHSLSHGAEKVAPPEVAPKCKSAAPKSKX
GAVKEGVNKSERKMLTLKGAADVDPDSCGLEHSAHVLEKGGKVSATLGLVDIVKGT
NSYIKQLLEDKDESRWTFIRSGRWGVTVIGSKLEHSAHVLEKGGKVSATLGLVDIVKGT
AHSKNFTYRPFYPLEIDYDDEAVKLTVPKTSKLPKQVLEIVGMFTFVESH
KVALVEYIDKMPGLKLSRQIOAAYSILSEVOAVSQGSESQIILDSNRYFTLI
PHDFGKKPPLNNADSVQAKVEMLDNLDIEVAYSLIRGSDSDSKPIDVNYEKLK
TDLKVYDRSEAEVIRKYKWNTHAYDLEVIDIPKIEREGESORYKPFROLHN
RLLWLSRRTNFAGLISGLRIATPEAYDGMFGKGIYFADVMKSKSNYCHTSQGD
PTGLILLGVALNGHVELKSHISKLPGKHSVKGKLTPTDPSASITLEGVEVPLG
TGIPSVNDTCLLYNEIYIVDYTAQVNLKLLKLFNFKTSLW"

CDS

BASE COUNT 1042 a 930 c 1080 g 793 t
ORIGIN

alignment_scores:

Quality: 518.50 Length: 156
Ratio: 3.812 Gaps: 1
Percent Similarity: 87.179 Percent Identity: 62.179

alignment_block:

US-09-236-995d-4 x BC012041

Align seg 1/1 to: BC012041 from: 1 to: 3845

1 AsnLysMetLeuLeuTrpHisGlySerArgLeuThrAsnPheValGly117
|||||
2625 AACGGAGGCTCTGTGGCAGCGCTCCAGGACCACCACTTGTGGCAT 2674
|||||
17 eLeuSerGlnGlyLeuArgIleAlaProProGluAlaProValThrGlyT 34
|||||
2675 CCTGTGCGAGGCTCTGGGATAGCCGCCACCTGAAGCGCTGTGACAGGCT 2724
|||||
34 yMetPheGlyLysGlyLeuTyrPheAlaAspLeuValSerLysSerAla 50
|||||
2725 ACATGTTGGGAAAGGATCTACTTTGGCGACATGGTTCACCAAGTGCA 2774
|||||
51 GlnTyrCysTyrValAspArgAsnAsnProValGlyLeuMetLeuLeuSe 67
|||||
2775 AACTACTGCCACACATCTCAGGAGACCGGATGGCTTAATACTGCTGGG 2824
|||||
67 rGluValAlaLeuGlyAspMetTyrGlnLeuLysLysAlaThrSerMetA 84
|||||
2825 AGAGGTGGCTTGGAAACATGTATGAACCTCAAGCATGCTTCACATATCA 2874
|||||
84 sPlysProArgGlyLysHisSerThrLysGlyLeuGlyLysThrVal 100
|||||
2875 GCAAGTTACCAAGGCAAGCAGTGTCAAAGGTTGGGAAAAACACC 2924
|||||
101 ProLeuGluSerGluPheValLysTrpArgAspValValValProCy 117
|||||
2925 CTGACCTTGGCCAGCATCACC...CTGAGGGGTGTAGAGTTCCTACT 2971
|||||
117 sGlyLysProValProSerSerIleArgSerSerGluLeuMetTyrAsnG 134
|||||
2972 GGAACAGGAGATCCCATCTGGTGTCAACGACACCTGCCTGCTGATAATG 3021
|||||
134 LuTyrIleValTyrAsnThrSerGlnValLysMetGlnPheLeuLeuLys 150
|||||
3022 AGTACATGTGTACGACATGCTCAGGTGAATCTCAATACCTGCTGAA 3071
|||||
151 ValArgPheHisLys 156
|||||
3072 CTCAGTCAATTTTAAG 3089

seq_name: gb_pr:HUMADPPO

seq_documentation_block:

LOCUS HUMADPPO 1771 bp mRNA PRI 30-OCT-1994
DEFINITION Human placental poly(ADP-ribose) polymerase mRNA, partial cds.

ACCESSION M17081

VERSION M17081.1 GI:178151

KEYWORDS

SOURCE Human placenta cDNA to mRNA, clone lambda-PAP803.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

TITLE

Molecular cloning of cDNA for human poly(ADP-ribose) polymerase and
expression of its gene during HL-60 cell differentiation [published
erratum appears in Biochem Biophys Res Commun 1987 Nov
13:148(3):1549]

JOURNAL Biochem. Biophys. Res. Commun. 146 (2), 403-409 (1987)

MEDLINE 87298455

REFERENCE 2 (bases 1321 to 1559)

AUTHORS

TITLE

JOURNAL

FEATURES

source

1..1771

/organism="Homo sapiens"

/db_xref="taxon:9606"

/map="1q41-q42"

Suzuki, H., Uchida, K., Shima, H., Sato, T., Okamoto, T., Kimura, T. and

Miwa, M.

Errata

Biochem.

Biophys. Res. Commun. 148, 1549-1550 (1987)

Location/Qualifiers

1..1771

/organism="Homo sapiens"

/db_xref="taxon:9606"

/map="1q41-q42"

```
gene 1. .1725
CDS /gene="ppol"
    /note="poly(ADP-ribose) polymerase"
    /codon_start=1
    /db_xref="GDB:G00-119-508"
    /protein_id="AAA51599.1"
    /db_xref="GI:178152"
    /translation="KKMEVEKEANIRVVSEDFLDYASASTKSLOELFLAHLSPWGAE
    VKAPVEVAPRGSGAALSKSGQVKEGINKSEKMKLTFLKGAADVDPDGLSHS
    AHVLEKSGKVFSAVLGLVDIVKGTNSYIKQLLEDKKNRYWIFRSGVRGTGVSNG
    LEQMPKEDANEHFMKLYBEKTNAMSKNFTYKPKFYPLEIDYQDDEAVKFLTVN
    PGTKSKLPKVPQILKMFIDVESMKRAMVEYIDLOKPLGLSKSQIOAAYSILSEV
    QOAVSGSSDSQIIDLNSRFTYLIPIHDFGKMKPPLNNADSVQAKVEMLDNLLDIEVA
    YSLLRGSSDSSKDPIDVNYEKLTKLVVDROSEAEIIRKYVKNTHATTHNAYOLE
    VDFIKERGEQORYKPKFLNRLHLLHGSRTNFAGILSGLRIAPPEAPVTGYM
    FGIQYFADWVSKSANYCHPTSQGDPIGLILLGEVALGNMYELKHASHISLPGKRHSV
    KGLGKTPDPFSAISLDGVDVPLGTGISSGVNDTSLLYNEYIYVDIAQVNLKYLK
    FNFKTSLW"

BASE COUNT 501 a 409 c 472 g 389 t
ORIGIN

alignment_scores:
  Quality: 517.50 Length: 156
  Ratio: 3.777 Gaps: 1
Percent Similarity: 87.821 Percent Identity: 62.179

alignment_block:
  US-09-236-995D-4 x HUMADPPO ..
  Align seg 1/1 to: HUMADPPO from: 1 to: 1771
    1 AsnLysMetLeuLeuTrpHisGlySerArgLeuThrAsnPheValGlyI1 17
    1246 AACCGAAGATTGCTGGCAGCGGTCCAGGACCACTTTGCTGGGAT 1295
    17 eLeuSerGlnGlyLeuArgIleAlaProProGluLeuAlaProValThrGlyT 34
    1296 CTGTGCCAGGCTCTCGGATAGCCCGCCGTCGAGCCGCGGTGACAGGCT 1345
    34 yMetPheGlyLysGlyLeuTyrrPheAlaAspLeuValSerLysSerAla 50
    1346 ACATGTTTGGTAAGGGATCTATTTCGCTGACATGCTCTCCAAAGAGTGCC 1395
    51 GlnTyrCysTyrValAspArgAsnAsnProValGlyLeuMetLeuLeuSe 67
    1396 AACTACTGCCATACGCTCTCAGGAGACCAATAGGCTTAATCCTGTTGGG 1445
    67 rGluValAlaLeuGlyAspMetTyrGluLeuLysLysAlaThrSerMetA 84
    1446 ACAAGTTCCTTGAACATGATGTAACCTGACGACGCTTCACATATCA 1495
    84 sPLysProProArgGlyLysHisSerThrLysGlyLysThrVal 100
    1496 GCAAGTTACCAAGGCAAGCACAGTGTCAAGGTTTGGGCAAACTACC 1545
    101 ProLeuGluSerGluPheValLysTrpArgAspValValProCy 117
    1546 CTGTATCTCTCAGCTAACATAGT...CTGATGCTGTAGACGTTCTCT 1592
    117 sGlyLysProValProSerSerIleArgSerSerGluLeuMetTyrAsnG 134
    1593 TGGGACCGGATTTCATCTGGTGTGAATGACACCTCTCTACTATATACG 1642
    134 luTyrIleValTyrAsnThrSerGlnValLysMetGlnPheLeuLys 150
    1643 AGTACATTGCTATGATATGCTCAGGTAAATCTGAAGTATCTGCTGAAA 1692
    151 ValArgPheHisHisLys 156
    1693 CTGAATTCATTTTAAG 1710
```

seq_name: gb_pr:HUMPOLP

seq_documentation_block:

LOCUS HUMPOLP 3640 bp mRNA 08-JAN-1995
DEFINITION Human poly(ADP-ribose) polymerase mRNA, complete cds.
ACCESSION M18112
VERSION M18112.1 GI:190166
KEYWORDS polymerase.
SOURCE Homo SV40 transformed fibroblast, cDNA to mRNA, clone pPAP.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

FEATURES

Source

1. .3640

/organism="Homo sapiens"

/db_xref="taxon:9606"

/map="1q41-q42"

<1. .3640

/note="POLP mRNA"

140. .3184

/gene="ppol"

140. .3184

/gene="ppol"

/note="poly(ADP-ribose) polymerase"

/codon_start=1

/db_xref="GDB:G00-119-508"

/protein_id="AAA60137.1"

/db_xref="GI:190167"

/translation="MAESSDKLYRVEYAKSGRASCCKSESIPKDSLRLMALWQSPMF

DGKVPWHYHSCFWKRGHSIRHPDVQVDFSELWDDQKVKRTAEAGVTGQGDGI

GSKAEKTLGDAFAEYAKSNRSTCRGCMKEIKGQVRLSKMVDPEKPOLGMDIRWYHP

GCFVKNREELGRPEYSASQLKGFSLATEDKALKKQLPGVSEGRKRGDGVDE

VAKKSKKEDKSKLEKALKAQNDLWIKNDKELKVCSTNDLKKELIFNKQVPSGE

SAILDRVADGVFGALILPCRECSQLVFKSDAYVCTGDTVAWTKCMVKTOTPNRKEV

TPKFERETSLYKLLKVKQDRIFFPETASVAATPPSTASAPAAVNSSASADKPLN

MKILTLGKLSRNKDEYKAMIEKLGKLTGTANKASLCISTKKEVERMKNKEEYKEAN

IRVVSDEFLQDVASATSLQLLEFLAHLSPWGAEVKAEPVEVAPRGSGAALSKSK

GVKBEKINKSEKRMKLTGKGAADVDPDGLSHSLEHLEKGGKVFSAVLGLDIKGT

NSYTKLQLEDKKNRYWIFRSGVRGTGVSNGKLEQMPKEDAEIHFMLYBEKTN

AWHSKNTFYKPKFYPLEIDYQDDEAVKLTVPNGTKSLPKPQDILKMFIDVESM

KKAMVEYIDLOKPLGLSKSQIOAAYSILSEVQOAVSOGSSDSOILDLNSRFTLI

PHDFGKMKPPLNNADSVQAKVEMLDNLLDIEVAYSLLRGSSDSSKDPIDVNYEKLK

RRLWHGSRRTNFAGILSGLRIAPPEAPVTGYMFGIYFADWVSKSANYCHPTSQD

PIGLITLGEVALGNMYELKHASHISLPGKRHSVKGKLTGTPDPFSAISLDGVDVPLG

TGISSGVNDTSLLYNEYIYVDIAQVNLKYLKFNFKTSLW"

BASE COUNT 999 a 833 c 1008 g 800 t

ORIGIN Chromosome 1p11-qter.

alignment_scores:

Quality: 517.50 Length: 156

Ratio: 3.777 Gaps: 1

Percent Similarity: 87.821 Percent Identity: 62.179

alignment_block:

US-09-236-995D-4 x HUMPOLP ..

Align seg 1/1 to: HUMPOLP from: 1 to: 3640

1 AsnLysMetLeuLeuTrpHisGlySerArgLeuThrAsnPheValGlyI1 17

||||: ||||||||||||||| |||||||:|||||

2705 AACCGAAGATTGCTGGCAGCGGTCCAGGACCACTTTGCTGGGAT 2754

17 eLeuSerGlnGlyLeuArgIleAlaProProGluAlaProValThrGlyT 34
|||||
2755 CTTGTCACAGGCTCTTCAGATAGCCGCCCTGAAGCGCCCTGCAGGCT 2804
|||||
34 yrMetPheGlyLysGlyLeuThrPheAlaAspLeuValSerLysSerAla 50
|||||
2805 ACATGTTGGTAAGGATGATTTTCGCTGACATGGTCTCCAAAGATGCC 2854
|||||
51 GlnTyrCysTyrValAspArgAsnProValGlyLeuMetLeuLeuSe 67
|||||
2855 AACTACTGCCATAGCTCTCAGGAGACCAATAGCTTAATCCTGTTGGG 2904
|||||
67 rGluValAlaLeuGlyAspMetTyrGluLeuLysAlaThrSerMetA 84
|||||
2905 AGAAGTTGCCCTTGGAACATGTTGAAGTGAAGGCTTTCACATATCA 2954
|||||
84 sPLysProProArgGlyLysHisSerThrLysGlyLeuGlyThrVal 100
|||||
2955 GCATGTTACCAAGGAGCAAGCAGATGTTCAAGGTTTGGGCAAACTACC 3004
|||||
101 ProLeuGluSerGluPheValLysTrpArgAspValValValProCy 117
|||||
3005 CCTGATCCTTCAGCTAATAGT...CTGGATGGTGTAGAGTTCTCTCT 3051
|||||
117 sGlyLysProProSerSerIleArgSerSerGluLeuMetTyrAsnG 134
|||||
3052 TGGACCGGGATTTTCATCGTGTGTAATGACACCTCTCTACTATATAACG 3101
|||||
134 luTyrIleValTyrAsnThrSerGlnValLysMetGlnPheLeuLeuLys 150
|||||
3102 AGTACATGTTCTATGATATTCCTCAGGTAAATCTGAAGTATCTGCTGAAA 3151
|||||
151 ValArgPheHisLysLys 156
|||||
3152 CTGAATTCATATTTAAG 3169
|||||

seq_name: gb_pat:A52134

seq_documentation_block:
LOCUS A52134 3792 bp DNA PAT 11-MAR-1997
DEFINITION Sequence 1 from Patent WO9618737.
ACCESSION A52134
VERSION A52134.1 GI:2304739
KEYWORDS
SOURCE unidentified
ORGANISM unidentified
unclassified.
REFERENCE
1 (bases 1 to 3792)
AUTHORS Buerkle,A., Zur,H.H. and Kuepper,J.
TITLE VECTORS AND VIRUSES FOR USE IN GENE THERAPY
JOURNAL Patent: WO 9618737-A 1 20-JUN-1996;
DEUTSCHES KREBSFORSCH (DE)
FEATURES
source Location/Qualifiers
1..3792
/organism="unidentified"
/db_xref="taxon:32644"
96..3137
/note="unnamed protein product"
/codon_start=1
/protein_id="CAA03385.1"
/db_xref="GI:2304740"
/translation="MAESDRLRYVEYAKSERASCKSEIPKDSLRMAIWQSPMF
DKGPHWHFSCFVGHSHIRHPDVEVDFGSELRWDDOKVKKTAAGGVTCGKGDI
GSAKTLGDFDAEYAKSNRSTCKGMEKIEGQVRLSKKMYDPEKPOLGMDLRVHP
GCFVKNRESELFRPEYSASQLKGFSLIATEDKEALKQLPGVKSEKRRKGDVGDVE
VAKKSKKEDKDLKLEKQNDLWNIDKELKVCSTNDLKELLIFNKQVPSGE
SAILDRVADGVFALLPCECGSLGVFKSDAYCTGDVTMTKCMVKVQTPTNRKEW
TPKEPTEISLKKLVKKODRIFPPETSASVAHPPPTASAPAAVNSASADKPLSKM
KLTGLKSLNRDEKVMKLEKLGKLTGTANKASLCISTKKKEKMKMKMEVKEANI
RVVSELDQVASTKSLQELFLAHLSPWAGAEVRAEPVEVVPAPKGAALSRSKSG
QYKEGINSKSRMKLITLKGGAADVPDGLSEHSAHVEKGVFATGLGLVDIVKGIN
SYVKLQLEDDEKFNRYWIPRSWGRVGTIGNSKLEKMPSEKDAIEHFMKLYBEKTGNA

WHSKFTYKPKFYPLEIDYCODEEAVKLTVPNTKSKLPKQVODLIKMLFDESMSK
KAMVEYEDLOKMPGLSKRQIOAAISILSEVOQVSSQSSQIIDLNSRPFYLLIP
HDFGKKPPLNNADSVQAKVEMLDNLLDIEVAYSLRGGSDSSKDPIDVNEKLLK
DLKVYDRSEAEIIRKYVKNTHATTHAYDLEVIDIFKIEREGECORYKPFQOLHNR
RLWHGSRRTTFAGILSGLRIAPPEAPVTGMEFGKGYFADMVSKSANTYCHTSQSDP
IGLILGELVALGNMVELKHASHISKLPKGHSVKGLGKTTTPDPSANISLDGVDVPLGT
GISSGVNDTSLLYNEYIVYDIAOVNLAYLLKLFNKFTSLW"

BASE COUNT 1048 a 846 c 1035 g 863 t

ORIGIN

alignment_scores:
Quality: 517.50 Length: 156
Ratio: 3.777 Gaps: 1
Percent Similarity: 87.821 Percent Identity: 62.179
alignment_block:
US-09-236-995D-4 x A52134
Align seg 1/1 to: A52134 from: 1 to: 3792

1 AsnLysMetLeuLeuThrHisGlySerArgLeuThrAsnPheValGlyTl 17
|||||
2658 AACCGAAGATTGCTGTGGCAGCGGTCCAGGACCAACCACTTTGCTGGAT 2707
|||||

17 eLeuSerGlnGlyLeuArgIleAlaProProGluAlaProValThrGlyT 34
|||||
2708 CTTGTCACAGGCTCTTCAGATAGCCGCCCTGAAGCGCCCTGCAGGCT 2757
|||||

34 yrMetPheGlyLysGlyLeuThrPheAlaAspLeuValSerLysSerAla 50
|||||
2758 ACATGTTGGTAAAGGATCTATTTCGCTGACATGCTTCCAAAGATGCC 2807
|||||

51 GlnTyrCysTyrValAspArgAsnProValGlyLeuMetLeuLeuSe 67
|||||
2808 AACTACTGCCATAGCTCTCAGGAGACCAATAGGCTTAATCCTGTTGG 2857
|||||

67 rGluValAlaLeuGlyAspMetTyrGluLeuLysLysAlaThrSerMetA 84
|||||
2858 AGAAGTTGCCCTTGGAACATGTTGAAGTGAAGCAGCTTTCACATATCA 2907
|||||

84 sPLysProProArgGlyLysHisSerThrLysGlyLeuGlyThrVal 100
|||||
2908 GCAAGTTACCAAGGCAAGCAGATGTTCAAGGTTTGGGCAAACTACC 2957
|||||

101 ProLeuGluSerGluPheValLysTrpArgAspValValValProCy 117
|||||
2958 CTTGATCCTTCAGCTAATAGT...CTGGATGGTGTAGAGCTTCTCT 3004
|||||

117 sGlyLysProProSerSerIleArgSerSerGluLeuMetTyrAsnG 134
|||||
3005 TGGACCGGGATTTTCATCTGGTGTGAATGACACCTCTCTACTATATAACG 3054
|||||

134 luTyrIleValTyrAsnThrSerGlnValLysMetGlnPheLeuLeuLys 150
|||||
3055 AGTACATGTTCTATGATATTCGCTCAGTAAATCTGAAGTATCTGCTGAAA 3104
|||||

151 ValArgPheHisLysLys 156
|||||3105 CTGAATTCATATTTAAG 3122
|||||

seq_name: gb_pr:HUMRISDAD

seq_documentation_block:
LOCUS HUMRISDAD 3795 bp mRNA PRI 09-JAN-1995
DEFINITION Human poly(ADP-ribose) synthetase mRNA, complete cds.
ACCESSION J03473
VERSION J03473.1 GI:337423
KEYWORDS poly(ADP-ribose) synthetase.
SOURCE Human placenta, cDNA to mRNA (library of H. Okayama), clones
PPARS[1,11,21,32,32,41,-F].
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;


```

/protein_id="AAF61293.1"
/db_xref="GI-7363322"
/translation="MKLTGGAADPDGSLHSAHVLEKGGKVFATLGLVDIVKGT
NSYKIQLEDDKESYWIIFRSRGVGTIGSNKLEQMPSEKDAVEHFMKLYEEKTGN
AWHSKDFTRYPKRYFLEIDYGDDEEAVKLTVPKGTSKLPVQELVGMIFDVESM
KRALVEIDLOKMPLEKLSRRIOQAAYSTLSEVQPPVSGSSESOLDLNSRFYTLI
PHDFGKKPPLNNADSLAKVEMLDNLLDIEVAYSLRGGSDSKDPIDVNYEKLK
TDIKVVDROSEAEVIRKYVKNTHATHNAYDLEVIDIEKIEREGESQRYKPRQLHN
RRLWGSRTTNFAGILSGLRIAPPEAPVTGTFMGKGIYFADVMVSKSANYCHTSGD
PGLIMGEVALGNMYELKHASHISKLPKHSVKGIGTKTTPDPSASITLEGVEVPLG
TGTPSVNDTCLLYNEYIYVDAQVNLKLLKFNFTSLW"
3'UTR
BASE COUNT 472 a 424 c 493 g 343 t
ORIGIN

```

```

alignment_scores:
  Quality: 516.50      Length: 156
  Ratio: 3.798        Gaps: 1
Percent Similarity: 87.179 Percent Identity: 61.538

```

```

alignment_block:
US-09-236-995D-4 x AFI26717

```

Align seg 1/1 to: AFI26717 from: 1 to: 1732

```

1 AsnLysMetLeuLeuTrpHisGlySerArgLeuThrAsnPheValGlyI1 17
|||||
1251 AACCGGAGGCTGCTGTGGCAGGCTCCAGGACCACCACTTTGTGGCAT 1300

17 eLeuSerGlnGlyLeuArgIleAlaProProGluAlaProValThrGlyT 34
|||||
1301 CCGTTCGCGAGGCTGCGGATAGCCACCTGAAGCGGCTGTGACGGCT 1350

34 yrMetPheGlyLysGlyLeuTrpPheAlaAspLeuValSerLysSerAla 50
|||||
1351 ACATGTTGGGAAGGATCTACTTTCGCGACATGGTGTCCAAAAGTGCA 1400

51 GlnTyrCysTyrValAspArgAsnProValGlyLeuMetLeuLeuSe 67
|||||
1401 AACTACTGCCACATCTCAGGAGACCGGATTGGCTTAATAATGCTGGG 1450

67 rGluValAlaLeuGlyAspMetTyrGluLeuLysLysAlaThrSerMetA 84
|||||
1451 AGAGTTCCTTGGAAATATGTATGAACCAAGCATGCTTCACATATCA 1500

84 spLysProProArgGlyLysHisSerThrLysGlyLeuGlyLysThrVal 100
|||||
1501 GCAGTACCAAGGCAAGGCAAGCAGTGTCAAGGTTGGGAAAACCAACC 1550

101 ProLeuGluSerGluPheValLysTrpArgAspValValValProCy 117
|||||
1551 CTGACCCCTTCGGCCAGCATCACC...CTGGAGGGTGTAGAGGTTCCACT 1597

117 sGlyLysProValProSerSerIleArgSerSerGluLeuMetTyrAsnG 134
|||||
1598 GGAACAGGATCCATCTGTGTGTCAACAGACACCTGCTGCTGTATATG 1647

134 luTyrIleValTyrAsnThrSerGlnValLysMetGlnPheLeuLeuLys 150
|||||
1648 AGTACATGTCTAGACATGCTCCTCAGGTGATCTCAAAATACCTGCTGAA 1697

151 ValArgPheHisLysLys 156
|||||
1698 CTCAGTTCATTTTAAG 1715

```

seq_name: gb_pl.AC006593

```

seq_documentation_block:
LOCUS AC006593 79663 bp DNA PLN 05-APR-2000
DEFINITION Arabidopsis thaliana chromosome II section 176 of 255 of the
complete sequence. Sequence from clones Fl6d14, T28P16.
ACCESSION AC006593 AE002093

```

AC006593.4 GI:6598616

HTG.

SOURCE

ORGANISM

Arabidopsis thaliana
thale cress.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE

AUTHORS

1 (bases 1 to 79663)
Lin,X., Kaul,S., Rounsley,S.D., Shea,T.P., Benito,M.-I., Town,C.D.,
Fuji,C.Y., Mason,T.M., Bowman,C.L., Barnstead,M.E.,
Feldblum,T.V., Buell,C.R., Ketchum,K.A., Lee,J.J., Ronning,C.M.,
Koo,H., Moffat,K.S., Cronin,L.A., Shen,M., VanAken,S.E., Unayam,L.,
Tallon,L.J., Gill,J.E., Adams,M.D., Carrera,A.J., Creasy,T.H.,
Goodman,H.M., Somerville,C.R., Copenhaver,G.P., Preuss,D.,
Nierman,W.C., White,O., Eisen,J.A., Salzberg,S.L., Fraser,C.M. and
Venter,J.C.

Sequence and analysis of chromosome 2 of the plant Arabidopsis

thaliana

Nature 402 (6763), 761-768 (1999)

JOURNAL

MEDLINE

PUBMED

REFERENCE

2 (bases 1 to 79663)

AUTHORS

TITLE

JOURNAL

COMMENT

Submitted (09-MAR-2000) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA
On Dec 17, 1999 this sequence version replaced gi:4432811.
The sequence and annotation of chromosome 2 were merged from those
of the individual clones on this chromosome after removing
overlaps. For detailed information, please see the TIGR web site
(http://www.tigr.org/tdb/at.html).

Genes were identified by a combination of three methods: Gene
prediction programs including GRAIL
(ftp://arthur.epm.ornl.gov/pub/xgrail), Genefinder (Phil Green,
University of Washington), Genscan (Chris Burge,
http://genomic.stanford.edu/GENSCAN.html), and NetPlantGene
(http://www.cbs.dtu.dk/services/NetGene2/), searches of the
complete sequence against a peptide database and plant EST
databases at TIGR, and manual curations based on those analyses.
Annotated genes are named to indicate the level of evidence for
their annotation. Genes with similarity to other proteins are named
after the database hits. Genes without significant peptide
similarity but with EST similarity are named as 'unknown' proteins.
Genes without protein or EST similarity, that are predicted by two
or more gene prediction programs over most of their length are
annotated as 'hypothetical' proteins. Genes encoding tRNAs are
predicted by tRNAscan-SE (Sean Eddy,
http://genome.wustl.edu/eddy/tRNAscan-SE/). Simple repeats were
identified by repeatmasker (Arian Smit,
http://ftp.genome.washington.edu/RM/RepeatMasker.html). Genes are
numbered from the top to bottom of the chromosome.

We thank the CSHL/WashU/ABI consortium for sequencing BAC clones
F6223, F5J6, T17A5, and T13L16, the ESSA group for sequencing clone
F13D4, and Scott Jackson, Jiming Jiang, Klaus Meyer, Eric Richards
and Satoshi Tabata for helpful assistance. In addition, we would
like to thank the TIGR Bioinformatics Department, especially Lixin
Zhou, Hanif Khalak, Michael E. Heaney, Lily Fu, Feng Liang, Jeremy
Peterson, Michael Holmes, and Delwood Richardson for software and
database support.

This work was supported by the National Science Foundation,
Department of Energy and the US Department of Agriculture.

Address all correspondence to: at@tigr.org.

FEATURES

source

```

1..79663
    location/Qualifiers
        /organism="Arabidopsis thaliana"
        /cultivar="Columbia"
        /db_xref="taxon:3702"
        /chromosome="II"
misc_feature
1..79462

```



```

34 ..... 34
73624 GTTCTCACTACTAGTTAACTGTTGAGAACTGGAACATTTTAAAC 73575
35 .....Met...PheGlyLysGlyLeuTyrPheAlaAspLeuValSerLy 48
3574 CCCATGTTTCAGTTTGGAAAGGATATACCTTGTGACCTTGTCAGTAA 73525
48 sSerAlaGlnTyrCysTyrValAspArgAsnAsnProValGlyLeuMetL 65
73524 RAGTCTCAGTACTGCTACACTGTTAAGAAAAATCCGGTGGTCTAATGC 73475
65 euLeuSerGluValAlaLeuGlyAspMetTyrGluLeuLysLysAlaThr 81
73474 TTCTGAGTGAAGTTGTCATTGGGAGAAATACATGAGCTAACAAAAAGCTAAG 73425
81 ..... 81
73424 GTGAGCCCATATATACAGTTTTCAGGTTATGCTACTAACCAATTTGGTTCT 73375
81 ..... 81
73374 TATCTTGAAACTATATGAACAACCTTAAGATATATACCTTCATATAT 73325
82 .....SerMetAspLysProProArg 88
73324 GACTTCTCTTATGTCCTCAATCTTATGATATATGGAATAAACCTCCGAGA 73275
89 GlyLysHisSerThrLysGlyLeuGlyLysThrValProLeuGluSerG1 105
73274 GGGAAACACTCGACCAAGAGGCTCGGCAAGAAAGTGCTCAAGATCCGA 73225
105 uPheValLysTrpArgAspValValValProCysGlyLysProValP 122
73224 GTTGCCAAAGTGGAGAGGTGATGCTGCTTCCCTGTGGAAACCTGTT 73175
122 roSerSerIleArgSerSerGluLeuMetTyrAsnGluTyrIleValTyr 138
73174 CATCAAGGTCAGGCTCTGAGCTTATGATACAAATGATATATGCTGCTAC 73125
139 AsnThrSerGlnVal.....LysMetGlnPheLeuLeu 149
73124 GATACAGCCAGGTAACCTGTGTTTCCAAATAATCTTCTCTGTT 73080

seq_name: gb_ov.XLPARPG

seq_documentation_block:
LOCUS XLPARPG 3617 bp mRNA VRT 18-JAN-1993
DEFINITION X.laevis PARP gene encoding poly(ADP-ribose) polymerase.
ACCESSION Z12139
VERSION Z12139.1 GI:64967
KEYWORDS poly(ADP-ribose) polymerase.
SOURCE African clawed frog.
ORGANISM
Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
Xenopodinae; Xenopus.
1 (bases 1 to 3617)
Saulier-le Drean,B.M.
Thesis (1992) Lab. de Biol. et Genet. du Developpement, Universite
de Rennes I. URA CNRS 256
2 (bases 1 to 3617)
Saulier-le Drean,B.M.
Direct Submission
Submitted (15-MAY-1992) Saulier-le Drean B.M., Lab. de Biol. et
Genet. du Developpement, Universite de Rennes I. URA CNRS 256,
Campus de Beaulieu, Av. du Gal Leclerc, Rennes-cedex, FRANCE, 35042
FEATURES
source
1..3617
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/germline
/sex="female"

```

```

/dev_stage="Adult"
/tissue_type="Ovary"
/cell_type="Oocyte"
/clone_lib="oligo dt primed lambda gt10 library of D.
MELTON"
/clone="clone parp 923"
1..2997
/gene="PARP"
1..2997
/gene="PARP"
/EC_number="2.4.2.30"
/function="poly ADP-riboseylation, DNA repair"
/Note="protein sequence is in conflict with the conceptual
translation: poly(ADP-ribose) polymerase"
/citation=[1]
/codon_start=1
/product="NAD(+) ADP-ribosyltransferase"
/protein_id="CAA78126.1"
/db_xref="GI:1334661"
/db_xref="SWISS-PROT:P31669"
/translation="AKSRASCKKGGDNIAKESGLAIWQSPMFDCGKVPWHYSCF
WKRAVLSQDIYGYTELWEDQEMIKAIETGAAAGAGDSKGGKGMTLNDFAAE
YAKNSACKCEQIEKQIRISKSVDPOLGMIDRWHPDCEVSGRELDLFLP
SYSASQLKQFTILSAEDKSLKLPVAKNEGRKADEVDGHSAAATKKIKKEKES
KLERLLKQETELIWHIDELKVKVCTNDLKELLIANKQOVPSGNTIVRVSDGAFG
ALLPCECSGQVFVKDAYCTGDSAWTKVAKTQTPNRKDWTPKEFHEIPLYKFF
KFKRIDRAFPCAAPTPLSPAPEPKTVEETPECKPLNTKYLILGLKSNKDEY
KTLLEGKGVAGSAHAKNICISTNKEVKKMSKMEVKANRVVSDDFLKEVESGK
SVQELSOFGISSWGAETKEAVQPTKQSPSSGVPAGSKGKVEKSGSKSKMKML
TVRGAAIDPDSELEDCHVETGKIFSATLGLVDITRTGTSYKILQLEHDDRSY
WFRSGRGTIVIGSKLEMSKDEAIEHFLNYQKTNWHSNFTKYPKFPYPL
EIDYQGEVGVKLSVAGTKSLAKPQBELIKLIDVESMKAMVFEIDLQKMPGL
KLSKROIQSAVSIISOVQAVSESLSEARLLDLSNOFTLPHDFGKMKPPLNNLEY
IQAKVQMLNDLLEAVAYSLRGAGDGEKDPIDVYKYEIKTDIKVVKADSEESRITC
DIVANHTADHINADLEVLFIKIDREGEYORYKPKOLHNRQLLWHGSESTTFAGIL
SQGLRIAPPEAPVTGYMFGKGYFADMSKSNYCHAMPGSPICLILGSEVGLGNHE
LKAASQITKLPKGVKSGKGTAPDPSATVQDGVDPVPLGKGTSANISDTSLLYNEY
IYVDIAQNLKYLKLFKNYKGGMMW"
3'UTR
polyA_signal 3551..3556
polyA_site 3574
BASE COUNT 1131 a 757 c 910 g 819 t
ORIGIN

alignment_scores:
Quality: 511.50 Length: 156
Ratio: 3.734 Gaps: 1
Percent Similarity: 87.821 Percent Identity: 62.179

alignment_block:
US-09-236-995D-4 x XLPARPG ..
Align seg 1/1 to: XLPARPG from: 1 to: 3617
1 AsnLysMetLeuLeuTrpHisGlySerArgLeuThrAsnPheValGlyLyl 17
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2515 AACCGCCAGCTGCTTTGGCAGCGCTCCGCAACCAATTTTGCAGGAAT 2564
17 eLeuSerGlnGlyLeuArgIleAlaProProGluAlaProValThrGlyT 34
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2565 ATTGTCTCAGGCTCTCCGAAATTCCTCCGCGAGAGCTCTCTGTACCGGT 2614
34 YrMetPheGlyLysGlyLeuTyrPheAlaAspLeuValSerLysSerAla 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2615 ATATGTTTGGCAAGAGTATCTATTTTGGCAGATGATGATCAAAAGTGCA 2664
51 GlnTyrCysTyrValAspArgAsnAsnProValGlyLeuMetLeuLeuSe 67
:|||||:|||||:|||||:|||||:|||||:|||||:
2665 AACTACTGTGTCAGCAATGCCGGGTAGCCCACTAGGCTGATCTTACTGG 2714
67 rGluValAlaLeuGlyAspMetTyrGluLeuLysLysAlaThrSerMetA 84
:|||||:|||||:|||||:|||||:|||||:|||||:

```

```
2715 GGAGGTGGCCCTCGGAACATGCATGAGCTGAAAGCAGCATCACAAATTA 2764
84 spLysProArgGlyLysHisSerThrLysGlyLeuGlyLysThrVal 100
   ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
2765 CAAAACTTCCCAAGGCAACACAGTGTAAAGGTTTGGGTAGAACTGCA 2814
101 ProLeuGluSerGluPheValLysTrpArgAspValValValProCy 117
   .. ||| ||| |||::: |||::: ||| ||| ||| ||| ||| |||
2815 CCAGATCCCTCAGCTACTGTACAG...CTGGATGGAGTGGATGTTCCCTCT 2861
117 sGlyLysProValProSerSerIleArgSerSerGluLeuMetTyrAsnG 134
   ||||| ::: :::: ||| :::: ||| :::: ||| :::: ||| ::::
2862 CGGAAAGGAACCTTCTGCAAAATATCAGTCACACACAGCCCTGTTGTATAATG 2911
134 luTyrIleValTyrAsnThrSerGlnValLysMetGlnPheLeuLeuLys 150
   ||||| ||||| |||::: :::: ||| :::: ||| :::: ||| ::::
2912 AATATATTGTGTATGATATTGCTCAGGTCAACCTGAAGTACCTGCTGAAG 2961
151 ValArgPheHisHisLys 156
   :::: ||| :::: |||
2962 CTCAGTTCAACTACAAA 2979
```

OM of: US-09-236-995D-4 to: N_Geneseq_1101.* out_format : pfs

Date: Mar 9, 2002 8:30 PM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

```
-MODEL=frame+p2n.model -DEV=xlp
-Q/cg2_1/USPTO.spool/US09236995/runat_08032002_113149_25135/app_query.fasta_1.1273
-DB=N_Geneseq_1101 -QMT=fastap -SUFFIX=mg -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-QCAPOP=4.500 -QCAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-DELOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM-ext -MINLEN=0 -MAXLEN=200000000
-USER=US09236995 -CGN1_1_309 -NCFU=6 -ICPU=3 -LONGLOG -NO_XLPXY
-WAIT -THREADS=1
```

Search information block:

```
Query: US-09-236-995D-4
Query length: 157
Database: N_Geneseq_1101.*
Database sequences: 930621
Database length: 428662619
Search time (sec): 250.160000
```

score_list:

Sequence	Strd Orig	2Score	EScore	Len	Documentation
/SIDS2/gcgdata/geneseq/NA1999.DAT:AA89542 +		828.00	1857.54	2949	2.2e-95
/SIDS2/gcgdata/geneseq/NA2000.DAT:AA260615 +		828.00	1856.60	3211	2.5e-95
/SIDS2/gcgdata/geneseq/NA2000.DAT:AA260618 +		828.00	1856.59	3212	2.5e-95
/SIDS2/gcgdata/geneseq/NA1999.DAT:AA223799 +		517.50	1144.26	3580	1.2e-55
/SIDS2/gcgdata/geneseq/NA1996.DAT:AA213732 +		517.50	1143.62	3793	1.3e-55
/SIDS2/gcgdata/geneseq/NA2001.DAT:AA263954 +		500.50	1107.11	3045	1.4e-53
/SIDS2/gcgdata/geneseq/NA2000.DAT:AA260617 +		496.50	1101.82	2147	2.7e-53
/SIDS2/gcgdata/geneseq/NA2000.DAT:AA260616 +		453.50	1002.60	2295	9.1e-48
/SIDS2/gcgdata/geneseq/NA2001.DAT:AA260615 +		441.50	985.91	864	7.7e-47
/SIDS2/gcgdata/geneseq/NA2001.DAT:AA260616 +		441.50	979.34	1566	1.8e-46
/SIDS2/gcgdata/geneseq/NA2001.DAT:AA260617 +		441.50	977.71	1814	2.2e-46
/SIDS2/gcgdata/geneseq/NA2001.DAT:AA260618 +		441.50	977.54	1843	2.3e-46
/SIDS2/gcgdata/geneseq/NA2001.DAT:AA260619 +		441.50	977.54	1843	2.3e-46
/SIDS2/gcgdata/geneseq/NA2001.DAT:AA260620 +		441.50	977.13	1912	2.4e-46
/SIDS2/gcgdata/geneseq/NA2001.DAT:AA260621 +		441.50	971.44	3200	4.9e-46
/SIDS2/gcgdata/geneseq/NA1995.DAT:AA260622 +		435.50	952.03	5345	5.9e-45
/SIDS2/gcgdata/geneseq/NA1995.DAT:AA260623 +		429.50	945.91	2682	1.3e-44
/SIDS2/gcgdata/geneseq/NA2001.DAT:AA260624 +		428.00	945.83	1980	1.3e-44
/SIDS2/gcgdata/geneseq/NA2000.DAT:AA260625 +		358.50	799.94	595	1.8e-36
/SIDS2/gcgdata/geneseq/NA1995.DAT:AA260626 +		350.50	770.74	1592	7.5e-35
/SIDS2/gcgdata/geneseq/NA2000.DAT:AA260627 +		227.50	489.07	1587	3.6e-19
/SIDS2/gcgdata/geneseq/NA2000.DAT:AA260628 +		227.50	488.05	1740	4.1e-19
/SIDS2/gcgdata/geneseq/NA2000.DAT:AA260629 +		225.50	480.56	2265	1.1e-18
/SIDS2/gcgdata/geneseq/NA2000.DAT:AA260630 +		225.50	480.56	2265	1.1e-18
/SIDS2/gcgdata/geneseq/NA2000.DAT:AA260631 +		173.00	350.53	5490	1.9e-11
/SIDS2/gcgdata/geneseq/NA2001.DAT:AA260632 +		173.00	350.53	5490	1.9e-11
/SIDS2/gcgdata/geneseq/NA2001.DAT:AA260633 +		133.50	287.19	472	6.4e-08
/SIDS2/gcgdata/geneseq/NA2001.DAT:AA260634 +		127.00	245.19	5482	1.4e-05
/SIDS2/gcgdata/geneseq/NA2000.DAT:AA260635 +		123.00	239.56	3984	2.9e-05
/SIDS2/gcgdata/geneseq/NA2000.DAT:AA260636 +		123.00	239.15	4134	3.0e-05
/SIDS2/gcgdata/geneseq/NA2000.DAT:AA260637 +		123.00	239.15	4134	3.0e-05
/SIDS2/gcgdata/geneseq/NA2000.DAT:AA260638 +		123.00	238.24	4491	3.4e-05
/SIDS2/gcgdata/geneseq/NA2001.DAT:AA260639 +		123.00	237.83	4657	3.6e-05
/SIDS2/gcgdata/geneseq/NA2001.DAT:AA260640 +		118.00	249.37	582	8.2e-06
/SIDS2/gcgdata/geneseq/NA2000.DAT:AA260641 +		118.00	240.12	1344	2.7e-05
/SIDS2/gcgdata/geneseq/NA2001.DAT:AA260642 +		118.00	231.35	2971	8.2e-05
/SIDS2/gcgdata/geneseq/NA2001.DAT:AA260643 +		118.00	230.35	3254	9.4e-05
/SIDS2/gcgdata/geneseq/NA2001.DAT:AA260644 +		118.00	230.16	3308	9.6e-05
/SIDS2/gcgdata/geneseq/NA2001.DAT:AA260645 +		118.00	230.02	3353	9.8e-05
/SIDS2/gcgdata/geneseq/NA1999.DAT:AA260646 +		118.00	229.86	3400	0.00010

```
/SIDS2/gcgdata/geneseq/NA2001.DAT:AA63837 + 118.00 229.52 0.0001 3508
/SIDS2/gcgdata/geneseq/NA2001.DAT:AA63838 + 118.00 228.64 0.0001 3797
/SIDS2/gcgdata/geneseq/NA2001.DAT:AA63918 + 118.00 228.63 0.0001 3799
/SIDS2/gcgdata/geneseq/NA2001.DAT:AA63918 + 118.00 228.59 0.0001 3816
/SIDS2/gcgdata/geneseq/NA2001.DAT:AA63926 + 118.00 227.72 0.0001 4127
```

seq_name: /SIDS2/gcgdata/geneseq/NA1999.DAT:AA89542

seq_documentation_block:

ID: AAX89542 standard; cDNA; 2949 BP.

AC: AAX89542;

DT: 06-OCT-1999 (first entry)

DE: Maize poly ADP-ribose polymerase gene.

KW: PARP; poly ADP-ribose polymerase; metabolism; maize; transform; ss.

OS: Zea mays.

FH: Key
FT: CDS
FT: Location/Qualifiers
FT: 1..2949

```
/*tag= a
/product= "PARP"
/feature= "Poly ADP-ribose polymerase"
/transl_except= (pos:1453..1455, aa:Xaa)
/transl_except= (pos:1495..1497, aa:Xaa)
/transl_except= (pos:1504..1506, aa:Xaa)
/transl_except= (pos:1534..1536, aa:Xaa)
/transl_except= (pos:1537..1539, aa:Xaa)
/transl_except= (pos:1561..1563, aa:Xaa)
/transl_except= (pos:1582..1584, aa:Xaa)
/transl_except= (pos:1588..1590, aa:Xaa)
/transl_except= (pos:1594..1596, aa:Xaa)
/transl_except= (pos:1975..1977, aa:Xaa)
/transl_except= (pos:2011..2013, aa:Xaa)
/transl_except= (pos:2014..2016, aa:Xaa)
/transl_except= (pos:2026..2028, aa:Xaa)
/transl_except= (pos:2047..2049, aa:Xaa)
/transl_except= (pos:2050..2052, aa:Xaa)
/transl_except= (pos:2053..2055, aa:Xaa)
/transl_except= (pos:2056..2058, aa:Xaa)
/transl_except= (pos:2077..2079, aa:Xaa)
/transl_except= (pos:2095..2097, aa:Xaa)
/transl_except= (pos:2107..2109, aa:Xaa)
/transl_except= (pos:2116..2118, aa:Xaa)
/transl_except= (pos:2119..2121, aa:Xaa)
/transl_except= (pos:2122..2124, aa:Xaa)
/transl_except= (pos:2125..2127, aa:Xaa)
/transl_except= (pos:2128..2130, aa:Xaa)
/feature= "Xaa = unknown"
```

WO9937789-A1.

29-JUL-1999.

26-JAN-1999; 99WO-US01591.

27-JAN-1998; 98US-0072785.

(PION-) PIONEER HI-BRED INT INC.

Mahajan P, Zuo Z;

WPI: 1999-444613/37.

P-PSDB; AAY28464.

New maize poly ADP-ribose polymerase gene useful in transforming plants to alter their metabolic state
Claim 1; Page 30-35; 48pp; English.

CC The sequence is the maize poly ADP-ribose polymerase (PARP) gene. The
CC sequence codes a PARP with 982 amino acids (AA28464). PARP is required
CC in the cell in most cases of DNA repair, recombination, rearrangement
CC and transposition. PARP gene and antisense gene can be used to transform
CC plant cells and alter the metabolic state of the transformed cell. This
CC is useful in enhancing disease resistance in plants and methods of
CC genetic transformation of plants. Plants transformed with either a sense
CC or antisense PARP nucleotide sequence may be utilized to increase
CC transformation frequency in plant cells. The enzyme also plays a role in
CC cellular stress, so may be beneficial for prevention of plant disease or
CC pathogen attack.
XX
SQ Sequence 2949 BP; 895 A; 539 C; 732 G; 756 T; 27 other;

alignment_scores:
Quality: 828.00 Length: 157
Ratio: 5.274 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-236-995D-4 x AAX89542 ..
Align seg 1/1 to: AAX89542 from: 1 to: 2949
1 AsnLysMetLeuLeuTrpHisGlySerArgLeuThrAsnPheValGlyI1 17
2476 AACACAGATGCTATTATGGCACGGTTCAAGGTTGACGAATTTTGTGGGAAT 2525
17 eLeuSerGlnGlyLeuArgIleAlaProProGluAlaProValThrGlyT 34
2526 TCTTAGTCAAGGGCTAAGAAATTTGCACCTCTCTGAGGCACCTGTACTGGCT 2575
34 yrMetPheGlyLysGlyLeuTyrPheAlaAspLeuValSerLysSerAla 50
2576 ATATGTTTCGCAAGGCCCTCTACTTTGCAGATCTAGTAAGCAAGAGCGCA 2625
51 GlnTyrCysTyrValAspArgAsnAsnProValGlyLeuMetLeuLeuSe 67
2626 CAATAGTGTATGTGGATAGGAATATCTGTAGGTTTGTAGTCTCTCTTC 2675
67 rGluValAlaLeuGlyAspMetTyrGluLeuLysLysAlaThrSerMetA 84
2676 TGAGGTTGCTTTAGGAGACATGTATGAACATAAGAAAGCCACGTCATGG 2725
84 sPlysProProArgGlyLysHisSerThrLysGlyLeuGlyLysThrVal 100
2726 ACAACCTCCAGAGGGAAGCATTCGACCAAGGGATTAGGCAAAACCGTG 2775
101 ProLeuGluSerGluPheValLysTyrArgAspValValValProcy 117
2776 CCACCTGGAGTCAGAGTTTGTGAAGTGGAGGGATGATGCTAGTTCCTCG 2825
117 sGlyLysProValProSerSerIleArgSerSerGluLeuMetTyrAsnG 134
2826 CGCAAGCCGGTGCCATCATCAATTAGGAGCTCTGAACATCATGTACAATG 2875
134 luTyrIleValTyrAsnThrSerGlnValLysMetGlnPheLeuLeuLys 150
2876 AGTACATCGTCTACACACATCCAGGTGAAGATGCAGTCTCTGTCTGAAG 2925
151 ValArgPheHisLysArg 157
2926 GTGCGTTTCATCCACAGAGG 2946

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AA260615
seq_documentation_block:
ID AA260615 standard; DNA; 3211 BP.
XX
AC AA260615;
XX
DT 16-MAY-2000 (first entry)

XX DE DNA encoding the poly(ADP-ribose) polymerase ZAP1 protein of Zea mays.
XX
XX ZAP1; poly(ADP-ribose) polymerase; PARP; poly(ADP-ribose) transferase;
KW programmed cell death; apoptosis; growth rate; stress; cold; pathogen;
KW pest; drought; heat; fungi; nematode; seed-shatter; ss.
XX
OS Zea mays.
XX
FH Key Location/Qualifiers
FT CDS 113..3022
FT /*tag= a
FT /product= "ZAP1 protein"
XX
PN WO200004173-A1.
XX
XX 27-JAN-2000.
PD
XX
PF 12-JUL-1999; 99WO-EP04940.
XX
PR 17-JUL-1998; 98US-0118276.
XX
XX (PLBZ) PLANT GENETIC SYSTEMS NV.
PA
XX Babychuk E, Kushnir S, De Block M;
PI
XX
XX WPI: 2000-182436/16.
DR P-PSDB; AAY68833.
XX
XX Modulating cell death, growth and stress resistance in eukaryotes,
PT specifically plants, used, e.g. to impart fungus or nematode resistance
PT
XX
PS Example 1; Page 79-84; 126pp; English.
XX
CC The present sequence encodes the ZAP1 protein of Zea mays. This protein
CC is a poly(ADP-ribose) polymerase (PARP) protein (also known as
CC poly(ADP-ribose) transferase). The PARP enzyme is involved in programmed
CC cell death or apoptosis, and is a nuclear enzyme. The ZAP1 polynucleotide
CC sequences can be used for modulation of programmed cell death in
CC eukaryotic cells. The method is used, specifically in plants, to induce,
CC or protect against, programmed cell death, depending on the extent to
CC which PARP activity is reduced. Reducing expression of endogenous ZAP
CC class PARP only is also used to modulate programmed cell death, to
CC increase growth rate and to produce plant cells that are more tolerant
CC of stress (cold, chemical treatments, pathogens, pests, drought, heat,
CC etc., or during transformation). Particular applications are generation
CC of plants that are resistant to fungi or nematodes; are male or female
CC sterile; or have better seed-shatter properties. The methods are also
CC used to improve growth of transformed plant cells (and derived calli or
CC complete plants).
XX
SQ Sequence 3211 BP; 968 A; 604 C; 813 G; 826 T; 0 other;

alignment_scores:
Quality: 828.00 Length: 157
Ratio: 5.274 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-236-995D-4 x AA260615 ..
Align seg 1/1 to: AA260615 from: 1 to: 3211
1 AsnLysMetLeuLeuTrpHisGlySerArgLeuThrAsnPheValGlyI1 17
2549 AACAGATGCTATTATGGCACGGTTCAAGGTTGACGAATTTTGTGGGAAT 2598
17 eLeuSerGlnGlyLeuArgIleAlaProProGluAlaProValThrGlyT 34
2599 TCTTAGTCAAGGGCTAAGAAATTTGCACCTCTCTGAGGCACCTGTACTGGCT 2648

34 yrmPheGlyGlyLeuTyrPheAlaAspLeuValSerLysSerAla 50
 |||||
 2649 ATATGTTTCGCAAGGCTCTACTTTCAGATCTAGTAAGCAAGAGCGCA 2698
 51 GlnTyrCysTyrValAspArgAsnProValGlyLeuMetLeuLeuSe 67
 |||||
 2699 CAATACGTGTTATGTGATAGGAATAATCCTGTAGGTTTGATGCTTCTTC 2748
 67 rGluValAlaLeuGlyAspMetTyrGluLeuLysLysAlaThrSerMetA 84
 |||||
 2749 TGAGTTGCTTTTAGGAGACATGTATGAACATAACAAAGCCAGCGTCCATGG 2798
 84 sPLysProProArgGlyLysHisSerThrLysGlyLeuGlyLysThrVal 100
 |||||
 2799 ACAAACTCCCAAGAGGGAAGCATTCGACCAAGGATTTAGCAAAACCGTG 2848
 101 ProLeuGluSerGluPheValLysTrpArgAspValValValProCy 117
 |||||
 2849 CCATGGAGTCAGAGTTTGGAAGTGGAGGATGATGCTGTAGTTCCTTG 2898
 117 sGlyLysProValProSerSerLysArgSerSerGluLeuMetTyrAsnG 134
 |||||
 2899 CGCAAGCGGTGCATCATCAATTAGGAGCTCTGAACCTCATGTACAAATG 2948
 134 LuTyrIleValTyrAsnThrSerGlnValLysMetGlnPheLeuLys 150
 |||||
 2949 AGTACATCGTCTACAAACACATCCAGGTGAAGATGCAGTTCCTTGTCTGAAG 2998
 151 ValArgPheHisLysLysArg 157
 |||||
 2999 GTGCGTTTCCATCAAGAGG 3019

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAZ60618

seq_documentation_block:
 ID AAZ60618 standard; DNA; 3212 BP.
 XX
 AC AAZ60618;
 XX
 DT 16-MAY-2000 (first entry)
 XX
 DE DNA encoding the poly(ADP-ribose) polymerase ZAP2 protein of Zea mays.
 XX
 KW ZAP2; poly(ADP-ribose) polymerase; PARP; poly(ADP-ribose) transferase;
 KW programmed cell death; apoptosis; growth rate; stress; cold; pathogen;
 KW pest; drought; heat; fungi; nematode; seed-shatter; ss.
 XX
 OS Zea mays.
 XX
 FH Key Location/Qualifiers
 FT CDS 81..3023
 FT /*tag= a
 FT /product= "ZAP2 protein"
 XX
 PN WO200004173-A1.
 XX
 XX
 PD 27-JAN-2000.
 XX
 PF 12-JUL-1999; 99WO-EP04940.
 XX
 PR 17-JUL-1998; 98US-0118276.
 XX
 XX (PLB2) PLANT GENETIC SYSTEMS NV.
 XX
 PI Babyiuchuk E, Kushnir S, De Block M;
 XX
 XX WPI; 2000-182436/16.
 DR P-PSDB; AAY68839.
 XX
 XX Modulating cell death, growth and stress resistance in eukaryotes,
 PT specifically plants, used, e.g. to impart fungus or nematode resistance
 PT
 XX

PS Disclosure; Page 103-108; 126pp; English.

XX The present sequence encodes the ZAP2 protein of Zea mays. This protein
 CC is a poly(ADP-ribose) polymerase (PARP) protein (also known as
 CC poly(ADP-ribose) transferase). The PARP enzyme is involved in programmed
 CC cell death or apoptosis, and is a nuclear enzyme. The ZAP2 polynucleotide
 CC sequences can be used for modulation of programmed cell death in
 CC eukaryotic cells. The method is used, specifically in plants, to induce,
 CC or protect against, programmed cell death, depending on the extent to
 CC which PARP activity is reduced. Reducing expression of endogenous ZAP
 CC class PARP only is also used to modulate programmed cell death, to
 CC increase growth rate and to produce plant cells that are more tolerant
 CC of stress (cold, chemical treatments, pathogens, pests, drought, heat,
 CC etc., or during transformation). Particular applications are generation
 CC of plants that are resistant to fungi or nematodes; are male or female
 CC sterile; or have better seed-shatter properties. The methods are also
 CC used to improve growth of transformed plant cells (and derived calli or
 CC complete plants).

SQ Sequence 3212 BP; 974 A; 600 C; 815 G; 823 T; 0 other;

alignment_scores:

Quality: 828.00 Length: 157
 Ratio: 5.274 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:..

US-09-236-995D-4 x AAZ60618 ..

Align seg 1/1 to: AAZ60618 from: 1 to: 3212

1 AsnLysMetLeuLeuTrpHisGlySerArgLeuThrAsnPheValGlyI 17
 |||||
 2550 ACAACAGTCTATTATGGCAGCGTTCAAGTTGACGAATTTGTGGGAAT 2599
 17 eLeuSerGlnGlyLeuArgIleAlaProProGluAlaProValThrGlyT 34
 |||||
 2600 TCTTAGTCAAGGGCTAAGAAATTGCACCTCTCTCAGGCACCTGTTACTGGCT 2649
 34 yrmPheGlyLysGlyLeuTyrPheAlaAspLeuValSerLysSerAla 50
 |||||
 2650 ATATGTTTCGCAAGGCTCTACTTTCAGATCTAGTAAGCAAGAGCGCA 2699
 51 GlnTyrCysTyrValAspArgAsnProValGlyLeuMetLeuLeuSe 67
 |||||
 2700 CAATACGTGTTATGTGATAGGAATAATCCTGTAGGTTTGATGCTTCTTC 2749
 67 rGluValAlaLeuGlyAspMetTyrGluLeuLysLysAlaThrSerMetA 84
 |||||
 2750 TGAGTTGCTTTTAGGAGACATGTATGAACATAACAAAGCCAGCGTCCATGG 2799
 84 sPLysProProArgGlyLysHisSerThrLysGlyLeuGlyLysThrVal 100
 |||||
 2800 ACAAACTCCCAAGAGGGAAGCATTCGACCAAGGATTTAGCAAAACCGTG 2849
 101 ProLeuGluSerGluPheValLysTrpArgAspValValValProCy 117
 |||||
 2850 CCATGGAGTCAGAGTTTGGAAGTGGAGGATGATGCTGTAGTTCCTTG 2899
 117 sGlyLysProValProSerSerLysArgSerSerGluLeuMetTyrAsnG 134
 |||||
 2900 CGCAAGCGGTGCATCATCAATTAGGAGCTCTGAACCTCATGTACAAATG 2949
 134 LuTyrIleValTyrAsnThrSerGlnValLysMetGlnPheLeuLys 150
 |||||
 2950 AGTACATCGTCTACAAACACATCCAGGTGAAGATGCAGTTCCTTGTCTGAAG 2999
 151 ValArgPheHisLysLysArg 157
 |||||
 3000 GTGCGTTTCCATCAAGAGG 3020

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAZ23799

DE DNA encoding the poly(ADP-ribose) polymerase NAP protein of Zea mays.
 XX NAP; poly(ADP-ribose) polymerase; PARP; poly(ADP-ribose) transferase;
 KW programmed cell death; apoptosis; growth rate; stress; cold; pathogen;
 KW pest; drought; heat; fungi; nematode; seed-shatter; ss.
 XX OS

XX Zea mays.

XX Key Location/Qualifiers
 XX CDS 107..2068
 XX FT /*tag= a
 XX FT /product= "NAP protein"

XX PN W0200004173-A1.

XX PD 27-JAN-2000.

XX PF 12-JUL-1999; 99WO-EP04940.

XX PR 17-JUL-1998; 98US-0118276.

XX PA (PLBZ) PLANT GENETIC SYSTEMS NV.

XX PI Babiychuk E, Kushnir S, De Block M;

XX DR WPI; 2000-182436/16.

XX DP P-PSDB; AAY68834.

XX Modulating cell death, growth and stress resistance in eukaryotes,
 PT specifically plants, used, e.g. to impart fungus or nematode resistance
 PT

XX Example 1; Page 89-92; 126pp; English.

XX The present sequence encodes the NAP protein of Zea mays. This protein
 CC is a poly(ADP-ribose) polymerase (PARP) protein (also known as
 CC poly(ADP-ribose) transferase). The PARP enzyme is involved in programmed
 CC cell death or apoptosis, and is a nuclear enzyme. The NAP polynucleotide
 CC sequences can be used for modulation of programmed cell death in
 CC eukaryotic cells. The method is used, specifically in plants, to induce,
 CC or protect against, programmed cell death, depending on the extent to
 CC which PARP activity is reduced. Reducing expression of endogenous NAP
 CC class PARP only is also used to modulate programmed cell death, to
 CC increase growth rate and to produce plant cells that are more tolerant
 CC of stress (cold, chemical treatments, pathogens, pests, drought, heat,
 CC etc., or during transformation). Particular applications are generation
 CC of plants that are resistant to fungi or nematodes; are male or female
 CC sterile; or have better seed-shatter properties. The methods are also
 CC used to improve growth of transformed plant cells (and derived calli or
 CC complete plants).

XX Sequence 2295 BP; 668 A; 447 C; 597 G; 583 T; 0 other;

alignment_scores:

Quality: 453.50 Length: 158
 Ratio: 3.599 Gaps: 1
 Percent Similarity: 79.747 Percent Identity: 57.595

alignment_block:

US-09-236-995D-4 x AA260616

Align seg 1/1 to: AA260616 from: 1 to: 2295

1 AsnLysMetLeuTyrPheHisGlySerArgLeuThrAsnPheValGlyI1 17
 |||||
 1589 AATAGGATGCTTTGTGGCATGTTCTCGGTGAGCAACTGGGCTGGGAT 1638

17 eLeuSerGlnGlyLeuArgIleAlaProGluAlaProValThrGlyT 34
 |||||

1639 CCTTCTCAGGCTCGCATCGCTCTCTCTGAGCACCTGTTACTGTT 1688

34 yrMetPheGlyLysGlyLeuTyrPheAlaAspLeuValSerLysSerAla 50

|||||
 1689 ACATGTTGGCAAGGGTGTCTTACTTGTGACATGTTTCAAGAGTGCA 1738
 51 GlnTyrCysTyrValAspArgAsnProValGlyLeuMetLeuLeu 67
 |||||
 1739 AACTATTGCTACGCCTCTGAAACATGATAGTCTGGAGTCTGCTTTATG 1788
 67 rGluValAlaLeuGlyAspMetTyrGluLeuLysLysAlaThr...SerM 83
 |||||
 1789 TGAGGTTCATTTGGCGCATATGATGAGCTACTGATGAGATGATGATG 1838
 83 etAspLysProProArgGlyLysHisSerThrLysGlyLeuGlyLysThr 99
 |||||
 1839 CTAATAACCTGCCCAAGGAAATTAAGATCCAAAGGAGTGGTCAACA 1888
 100 ValProLeuGluSerGluPheValLysTyrArgAspValValVal 116
 |||||
 1889 GCACCTAACATGTCGAGTCTAAGTGTGCTGAGTGTGTTGTTCTCC 1938
 116 oCysGlyLysProValProSerSerIleArgSerSerGluLeuMetTyrA 133
 |||||
 1939 CCTTGGCGAACCCAAACAGAACCTTCCAAAGGGGTGCTGCTTTATA 1988
 133 snGluTyrIleValTyrAsnThrSerGlnValLysMetGlnPheLeuLeu 149
 |||||
 1989 ATGAGTACATAGTGTACAACTAGACACAGATAAGATGCGGTATGCTTA 2038
 150 LysValArgPheHisLysArg 157
 |||||
 2039 CATGTTAACTTCAATTTCAAGAGA 2062

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT: AAC85320

seq_documentation_block:

ID AAC85320 standard; cDNA; 864 BP.
 XX
 AC AAC85320;
 XX
 DT 29-MAR-2001 (first entry)
 XX
 DE 3' end fragment of hparp2, 3'-hPARP2.

XX Human; poly(ADP-ribose) polymerase; hPARP2; oxidative stress; ARDS;
 KW inflammation; ischaemic stroke; hemorrhagic shock; myocardial ischemia;
 KW infarction; cerebral vasospasm; rheumatoid arthritis; osteoarthritis;
 KW gouty arthritis; spondylitis; Behcet's disease; sepsis; septic shock;
 KW endotoxic shock; gram negative sepsis; gram positive sepsis; trauma;
 KW toxic shock syndrome; multiple organ injury syndrome; vasculitis;
 KW hemorrhage; conjunctivitis; uveitis; thyroid-associated ophthalmopathy;
 KW eosinophilic granuloma; asthma; chronic bronchitis; allergic rhinitis;
 KW pulmonary sarcoidosis; pleurisy; alveolitis; pneumonia; myocardium;
 KW bronchiectasis; pulmonary oxygen toxicity; keloid formation; brain;
 KW scar tissue formation; atherosclerosis; systemic lupus erythematosus;
 KW autoimmune thyroiditis; multiple sclerosis; Reynaud's syndrome;
 KW graft versus host disease; allograft rejection; cystic fibrosis;
 KW chronic glomerulonephritis; inflammatory bowel disease;
 KW Crohn's disease; ulcerative colitis; necrotizing enterocolitis;
 KW inflammatory dermatosis; contact dermatitis; atopic dermatitis;
 KW psoriasis; urticaria; fever; myalgia; meningitis; encephalitis;
 KW Sjogren's syndrome; alcoholic hepatitis; bacterial pneumonia;
 KW hypovolemic shock; Type 1 diabetes mellitus; hypersensitivity;
 KW leukocyte dyscrasia; thermal injury; cytokine-induced toxicity;
 KW expressed sequence tag; EST; RACE; PCR; amplify; primer;
 KW polymerase chain reaction; ds.

XX Synthetic.

XX Key Location/Qualifiers
 XX PH 1..864
 XX CDS /*tag= a
 FT /product= "hPARP2 C-terminal fragment"
 FT
 XX

CC neurodegenerative diseases (e.g., Parkinson's disease and Alzheimer's
 CC disease). shbPARS2 nucleic acids are useful as diagnostic reagents for
 CC detecting mutations in the associated gene; as hybridisation probes
 CC to isolate full-length shbPARS2 cDNAs and shbPARS2 genomic clones; and
 CC for chromosome localisation studies. The shbPARS2 protein is also useful
 CC as an immunogen to produce antibodies for therapeutic use. shbPARS2
 CC proteins, nucleotides and antibodies are also useful in screening methods
 CC for detecting the effect of added compounds on the production of mRNA and
 CC protein in cells.

XX
 SQ Sequence 1566 BP; 490 A; 328 C; 381 G; 367 T; 0 other;

alignment_scores:
 Quality: 441.50 Length: 158
 Ratio: 3.422 Gaps: 3
 Percent Similarity: 81.646 Percent Identity: 57.595

alignment_block:
 US-09-236-995D-4 x AAF59996 ..

Align seg 1/1 to: AAF59996 from: 1 to: 1566

```

1 AsnLysMetLeuLeuTrpHisGlySerArgLeuThrAsnPheValGlyI 17
|||||
1078 AACAGGATGCTTCTATGGCATGGTTCAGGATGAGTAACATGGGTGGGAAT 1127
|||||
17 eLeuSerGlnGlyLeuArgIleAlaProGluAlaProValThrGlyT 34
|||||
1128 CTGAGCCATGGGCTTCGAATGGCCACCTGAAGCTCCCATCACAGTT 1177
|||||
34 yrMetPheGlyGlyGlyLeuTyrPheAlaAspLeuValSerLysSerAla 50
|||||
1178 ACATGTTGGGAAGGAATCTACITGCTGACATGCTCTCCCAAGAGTGCC 1227
|||||
51 GlnTyrCysTyrValAspArgAsnProValGlyLeuMetLeuLeuSe 67
|||||
1228 AATTACTGCTTGGCTCTCGCTTAAGAATAACAGGACTGCTCTTATC 1277
|||||
67 rGluValAlaLeuLeuGlyAspMetTyrGluLeuLysLysAlaThr...SerM 83
|||||
1278 AGAGTACTCTAGTCTAGTCTATGACTACTAGAGGCCAATCTTAAAG 1327
|||||
83 etAspLysProProArgGlyLysHisSerThrLysGlyLeuGlyLysThr 99
|||||
1328 CCGAAGGATTCCTTCAAGGTAACATACACCAAGGGCTGGCAAGATG 1377
|||||
100 ValProLeuGluSerGluPheValLysTrpArgAspValValPr 116
|||||
1378 GCTCCAGTTCTGCTCCACTGCTGACCCCTGAATGGAGT...ACAGTGCC 1424
|||||
116 oCysGlyLysProValProSerSerIleArgSerSerGlu.....L 130
|||||
1425 ATTAGACACAGCAGTGCACAGGAATTCCTGAATCCAGATGTTATACCC 1474
|||||
130 euMetTyrAsnGluTyrIleValTyrAsnThrSerGlnValLysMetGln 146
|||||
1475 TCACCTACAAATGAATATATTATATATACCCCAACACAGGTCGGTATCGG 1524
|||||
147 PheLeuLeuLysValArgPheHis 154
|||||
1525 TACCTTTTAAAGGTTCACTTTTAAT 1548

```

seq_name: /SDS2/gcdata/geneseq/geneseq/NA2001.DAT.AAC85303

seq_documentation_block:

ID AAC85303 standard; cDNA; 1814 BP.

XX

AC AAC85303;

XX

DT 29-MAR-2001 (first entry)

XX

DE hparp2 cDNA.

XX Human; poly(ADP-ribose) polymerase; hPARP2; oxidative stress; ARDS;
 KW inflammation; ischaemic stroke; hemorrhagic shock; myocardial ischemia;
 KW infarction; cerebral vasospasm; rheumatoid arthritis; osteoarthritis;
 KW gouty arthritis; spondylitis; Behcet's disease; sepsis; septic shock;
 KW endotoxic shock; gram negative sepsis; gram positive sepsis; trauma;
 KW toxic shock syndrome; multiple organ injury syndrome; vasculitis;
 KW hemorrhage; conjunctivitis; uveitis; thyroid-associated ophthalmopathy;
 KW eosinophilic granuloma; asthma; chronic bronchitis; allergic rhinitis;
 KW chronic obstructive pulmonary disease; silicosis; reperfusion injury;
 KW pulmonary sarcoidosis; pleurisy; alveolitis; pneumonia; myocardium;
 KW bronchiectasis; pulmonary oxygen toxicity; keloid formation; brain;
 KW scar tissue formation; atherosclerosis; systemic lupus erythematosus;
 KW autoimmune thyroiditis; multiple sclerosis; Reynaud's syndrome;
 KW graft versus host disease; allograft rejection; cystic fibrosis;
 KW chronic glomerulonephritis; inflammatory bowel disease;
 KW Crohn's disease; ulcerative colitis; necrotizing enterocolitis;
 KW inflammatory dermatosis; contact dermatitis; atopic dermatitis;
 KW psoriasis; urticaria; fever; myalgia; meningitis; encephalitis;
 KW Sjogren's syndrome; alcoholic hepatitis; bacterial pneumonia;
 KW hypovolemic shock; Type 1 diabetes mellitus; hypersensitivity;
 KW leukocyte dyscrasia; thermal injury; cytokine-induced toxicity; ds.

XX Homo sapiens.

OS
 XX
 XX Key Location/Qualifiers
 CDS 63..1814
 FT /*tag= a
 FT /product= "hPARP2"
 XX
 PN WO2000771179-A2.
 XX
 XX 21-DEC-2000.
 PD
 PE 16-JUN-2000; 2000WO-US16629.
 XX
 PR 16-JUN-1999; 99US-0139543.
 XX
 PA (ICOS-) ICOS CORP.
 XX
 PI Christenson E, Demaggio AJ, Goldman PS, McElligott DL;
 XX P-PSDB; AAB47029.
 DR WPI; 2001-025335/03.
 XX P-PSDB; AAB47029.
 PT New human poly(ADP-ribose) polymerase for treating inflammatory,
 PT neurological, cardiovascular, or neoplastic tissue growth disorders,
 PT such as, arthritis, encephalitis, myocardial ischemia, and leukocyte
 PT metastasis

Claim 1; Page 91-93; 129pp; English.

XX This sequence encodes human poly(ADP-ribose) polymerase (hPARP2).
 CC This protein causes the covalent addition of polymers of ADP-ribose
 CC to protein targets. hPARP2 activity is induced in many instances of
 CC oxidative stress or during inflammation where there is direct damage
 CC to the DNA. hPARP2 may be used to identify antagonists which
 CC may be used to treat a human having a disorder mediated by PARP2
 CC activity, such as, inflammatory, neurological, cardiovascular,
 CC or neoplastic tissue growth disorders, e.g. ischemic stroke,
 CC hemorrhagic shock, myocardial ischemia or infarction,
 CC transplantation, cerebral vasospasm; rheumatoid, osteo- or gouty
 CC arthritis, spondylitis; Behcet's disease; sepsis, septic or endotoxic
 CC shock, gram negative or positive sepsis, toxic shock syndrome; multiple
 CC organ injury syndrome secondary to septicemia, trauma, or hemorrhage;
 CC allergic or vernal conjunctivitis, uveitis, thyroid-associated
 CC ophthalmopathy; eosinophilic granuloma; asthma, chronic bronchitis,
 CC allergic rhinitis, ARDS, chronic obstructive pulmonary disease,
 CC silicosis, pulmonary sarcoidosis, pleurisy, alveolitis, vasculitis,
 CC pneumonia, bronchiectasis, pulmonary oxygen toxicity, reperfusion
 CC injury of the myocardium, brain or extremities; cystic fibrosis; keloid
 CC formation, scar tissue formation; atherosclerosis; systemic lupus
 CC erythematosus, autoimmune thyroiditis, multiple sclerosis; Reynaud's

CC syndrome; graft versus host disease, allograft rejection; chronic
CC glomerulonephritis; inflammatory bowel disease, Crohn's disease,
CC ulcerative colitis; necrotizing enterocolitis; inflammatory dermatoses,
CC contact or atopic dermatitis, psoriasis, urticaria, fever and myalgias
CC due to infection; meningitis, encephalitis, and brain and spinal cord
CC injury due to minor trauma; Sjogren's syndrome; diseases involving
CC leukocyte diapedesis; alcoholic hepatitis; bacterial pneumonia;
CC antigen-antibody complex mediated diseases; hypovolemic shock; Type 1
CC diabetes mellitus; acute and delayed hypersensitivity; disease states
CC due to leukocyte dyscrasia and metastasis; thermal injury; granulocyte
CC transfusion associated syndromes; and cytokine-induced toxicity.
CC hPARP2 and antibodies to it, can also be used to diagnose these
CC conditions.
XX
SQ Sequence 1814 BP; 558 A; 389 C; 462 G; 405 T; 0 other;

alignment_scores:
Quality: 441.50 Length: 158
Ratio: 3.422 Gaps: 3
Percent Similarity: 81.646 Percent Identity: 57.595

alignment_block:

US-09-236-995D-4 x AAC85303 ..

Align seg 1/1 to: AAC85303 from: 1 to: 1814

1 AsnLysMetLeuLeuTrpHisGlySerArgLeuThrAsnPheValGlyI 17
1326 AACAGGATGCTTCTATGGCATGGTTCAGGATGAGTAACCTGGTGGGAAT 1375
17 eLeuSerGlnGlyLeuArgIleAlaProGluAlaProValThrGlyT 34
1376 CTTGAGCCATGGCTTCGAATTCGCCACCTGAAGCTCCCATCACAGTT 1425
34 yMetPheGlyLeuGlyLeuTyrPheAlaAspLeuValSerLysSerAla 50
1426 ACATGTTTGGGAAGGAATCTACTTGTGCATGTCTTCCAGAGTGC 1475
51 GlnTyrCysTyrValAspArgAsnAsnProValGlyLeuMetLeuLeu 67
1476 AATTACTGCTTTGGCTCTCGCCTAAGAAATACAGGACTGCTCTTATC 1525
67 rGluValAlaLeuGlyAspMetTyrGluLeuLysLysAlaThr...SerM 83
1526 AGAGTAGTCTAGTCTAGTCTAGTCTAGTCTAGTCTAGTCTAGTCTAG 1575
83 eAspLysProProArgGlyLysHisSerThrLysGlyLeuGlyLysThr 99
1576 CCGAAGGATGCTTCAAGGTAAACATACCAAGGGCTGGGCAAGATG 1625
100 ValProLeuGluSerGluPheValLysTrpArgAspValValValPr 116
1626 GCTCCAGTCTGCGCCATCTGTCACCTGAATGGAGT...ACAGTGCC 1672
116 oCysGlyLysProValProSerSerIleArgSerSerGlu.....L 130
1673 ATTAGGACACGCAAGTGCACAGGAATCTCAATCCAGATGTTATACCC 1722
130 eMetTyrAsnGluTyrIleValTyrAsnThrSerGlnValLysMetGln 146
1723 TCAACTACAAATGAATATATTGTATATAACCAACCAAGGTCCTGATCGG 1772
147 PheLeuLeuLysValArgPheHis 154
1773 TACCTTTTAAGGTTCAAGTTAAT 1796

seq_name: /SIDS2/gcdata/geneseq/geneseqn/NA2000.DAT.AA244287

seq_documentation_block:

ID_AA244287 standard; cDNA; 1843 BP.

XX

AC AA244287;

XX 31-MAR-2000 (first entry)
XX Human brain PARP2 cDNA.
DE
XX PARP; poly(ADP-ribose) polymerase; human; murine; detection; sepsis;
KW diagnosis; therapy; necrosis; apoptosis; neurodegenerative illness;
KW ischemic tissue damage; ss.
OS
XX Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 3..1715
FT /*tag= a
FT /product= "PARP2"
XX
XX WO9964572-A2.
XX 16-DEC-1999.
XX
XX 04-JUN-1999; 99WO-EP03889.
XX
XX 05-JUN-1998; 98DE-1025213.
XX 01-MAR-1999; 99DE-1008837.
XX
XX (BADI) BASF AG.
XX Kock M, Hoeger T, Kroeger B, Otterbach B, Lubisch W, Lemaire H;
XX WPI; 2000-087218/07.
XX P-PSDB; AAY51174.
XX
XX Novel genes and proteins, antibodies and binding partners useful in
XX diagnosis and therapy of energy deficiency associated disease
XX conditions
XX
XX Claim 7a; Page 49-52; 96pp; German.
XX
XX This invention describes novel human and murine poly(ADP-ribose)
XX polymerase (PARP) homologues, which are characterised by an amino acid
XX sequence with a functional NAD⁺-binding site and no zinc finger
XX sequence motif, of general formula CX₂CX₂2MHX₂2C (1). The nucleic acid
XX sequences, PARP homologues and antibodies are useful for analytic
XX detection of PARP homologues and for identifying PARP effectors or
XX binding partners, as well as for determining their effectiveness.
XX PARP-binding partners are useful for the diagnosis or therapy of a
XX disease condition, which is the result of a PARP protein, especially an
XX energy deficiency, which may comprise tissue damage from cell death
XX following necrosis or apoptosis. The disease condition may be chosen
XX from a neurodegenerative illness, or sepsis or ischemic tissue damage,
XX in particular neurotoxic disturbances, etc. This sequence encodes the
XX human PARP2 protein used in the method of the invention.
XX
SQ Sequence 1843 BP; 599 A; 377 C; 447 G; 420 T; 0 other;

alignment_scores:

Quality: 441.50 Length: 158
Ratio: 3.422 Gaps: 3
Percent Similarity: 81.646 Percent Identity: 57.595

alignment_block:

US-09-236-995D-4 x AA244287 ..

Align seg 1/1 to: AA244287 from: 1 to: 1843

1 AsnLysMetLeuLeuTrpHisGlySerArgLeuThrAsnPheValGlyI 17
1227 AACAGGATGCTTCTATGGCATGGTTCAGGATGAGTAACCTGGTGGGAAT 1276
17 eLeuSerGlnGlyLeuArgIleAlaProGluAlaProValThrGlyT 34
1277 CTTGAGCCATGGCTTCTGAATTCGCCACCTGAAGCTCCCATCACAGTT 1326

This invention describes novel 4-substituted 2H-phthalazin-1-one derivatives (I) which are used for the treatment or prophylaxis of diseases associated with elevated poly-(ADP-ribose)-polymerase (PARP; EC 2.4.2.30) activity. The products of the invention have neurotropic, neuroprotective, cerebroprotective, antiparkinsonian, nephroprotective, cardiatic, vasotropic, anticonvulsant, cytostatic, antibacterial, immunosuppressive, antiinflammatory, antirheumatic, antiarthritic, antidiabetic. (I) are especially used for treating or preventing neurodegenerative disease or neuronal damage (specifically associated with ischemia, trauma or massive bleeding, especially apoplexy or spinal-cranial trauma; or Alzheimer's disease, Huntington's disease or Parkinson's disease), treating or preventing ischemic damage (specifically renal damage after renal ischemia or during and after kidney transplantation or heart damage after cardiac ischemia), treating epilepsy, specifically generalized epileptic attacks (e.g. petit mal and tonic-clonic attacks) or partial epileptic attacks (e.g. temporal lobe and complex partial attacks), treating microinfarction (e.g. during and after heart valve replacement, aneurysm resectioning and heart transplantation), revascularization of critically constricted coronary arteries (e.g. after PCTA or by-pass operations) or peripheral arteries (e.g. leg arteries), treating acute myocardial ischemia and damage during or after its mechanical or drug-induced lysis and treating tumors and their metastasis, sepsis and septic shock, inflammatory and rheumatic disease (e.g. rheumatoid arthritis), and diabetes mellitus. Although (I) inhibit PARP (i.e. the known form designated PARP1), they especially selectively and strongly inhibit PARP homologs, specifically the homolog PARP2. In particular (I) have very strong PARP2 inhibitory activity (e.g. with K_i values of 1-20 nM) and high selectivity for PARP2 relative to PARP1 (generally by a factor of more than 5).

sequence 1843 BP: 599 A: 377 C: 447 G: 420 T: 0 other;

alignment_scores:	Quality:	441.50	Length:	158
-------------------	----------	--------	---------	-----

Percent	Ratio:	3.422	Gaps:	3
similarity:		81	Identity:	57.595

```

gn seg 1/1 to: AAC82090 from: 1 to: 1843
1 AsnLysMetLeuLeuTrpHisGlySerArgLeuThrAsnPheValGlyI 17
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1227 AACAGGATGCTTATGGCATGTTCCAGATGAGTAACATGGGTGGGAAT 1276
17 eLeuSerGlnClyLeuArgIleAlaProProGluAlaProValThrGlyT 34
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1326

```

OS Homo sapiens.

[illegible]

Environ Monit Assess (2017) 189:67

1426

[illegible][illegible]

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

[illegible]

T47 / CC9AAGG8110CCTCATACCGATCCGTCGGTGT

100 ValProLeuGluSerGluPheValLysTrpArgAspValValValPr ILI
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
1527 GCTCCCAAGTCTGCCCACTTCGTACCCCTGAATGGAGT...ACAGTGCC 1573

116 cCysGlyLysProValProSerSerIleArgSer:SerGlu.....L 130
 1574 ATTAGGACCAGCAAGTGCACACAGGAATTCGAATCCACATGGTTATACCC 1623
 130 eumetTyrAsnGluTyrIleValTyrAsnThrSerGlnValLysMetGln 146
 1624 TCACTACAATGAATATATTGATATATACCCCAACCAAGGTCGATGCGG 1673
 147 PheLeuLeuLysValArgPheHis 154
 1674 TACCTTTAAAGGTTCAAGTTTAAT 1697

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT: AAC77118

seq_documentation_block:

ID AAC77118 standard; cDNA; 1912 BP.

AC AAC77118;

DT 08-FEB-2001 (first entry)

XX Human ORFX ORF2673 polynucleotide sequence SEQ ID NO:5345.

XX Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
 KW vulnery; antipariatic; antiparkinsonian; nootropic; neuroprotective;
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
 KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
 KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
 KW thrombosis; contraceptive; ss.

XX Homo sapiens.

XX WO200058473-A2.

XX 05-OCT-2000.

XX 31-MAR-2000; 2000WO-US08621.

XX 31-MAR-1999; 99US-0127607.

XX 02-APR-1999; 99US-0127636.

XX 05-APR-1999; 99US-0127728.

XX 30-MAR-2000; 2000US-0540763.

XX (CURA-) CURAGEN CORP.

XX Shinkets RA, Leach M;

XX WPI; 2000-602362/57.

XX P-PSDB; AAB42909.

XX Novel nucleic acids and peptides derived from open reading frame X,

XX useful for treating e.g. cancers, proliferative disorders,

XX neurodegenerative disorders and cardiovascular disease -

XX Claim 5; Page 4521-4532; 5507pp; English.

XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatotropic; vulnery;
 CC antipariatic; antiparkinsonian; nootropic; neuroprotective;
 CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
 CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;
 CC antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;

CC antithyroid; and antianemic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORFX-associated disorder. The
 CC nucleic acids can be used to express ORFX proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.

XX Sequence 1912 BP; 620 A; 390 C; 467 G; 433 T; 2 other;

alignment_scores:

Quality: 441.50 Length: 158

Ratio: 3.422 Gaps: 3

Percent Similarity: 81.646 Percent Identity: 57.595

alignment_block:

US-09-236-995D-4 x AAC77118 ..

Align seg 1/1 to: AAC77118 from: 1 to: 1912

1 AsnLysMetLeuLeuTyrPheHisGlySerArgLeuThrAsnPheValGlyI 17
 1291 AACAGGATGCTTCTATGTCATGGTTCAGGATGAGTAAGTGGTGGGAAT 1340
 17 eLeuSerGlnGlyLeuArgIleAlaProGluAlaProValThrGlyT 34
 1341 CTTGAGCCATGGCTTCGAATGGCCACCTGAAGCTCCCATCAGGTT 1390
 34 yrMetPheGlyGlyGlyLeuTyrPheAlaAspLeuValSerLysSerAla 50
 1391 ACATGTTTGGGAAGGAATCTACTTTGCTGACATGCTTCCACAGTGCC 1440
 51 GlnTyrCysTyrValAspArgAsnAsnProValGlyLeuMetLeuLeuSe 67
 1441 AATTACTGCTTTCCTCTCGCCTAAAGAAATACAGGACTGCTGCTTTATC 1490
 67 rGluValAlaLeuGlyAspMetTyrGluLeuLysLysAlaThr...SerM 83
 1491 AGAGTAGCTCTAGTCAAGTGAATGAAGTACTAGAGGCCAATCCTAAGG 1540
 83 etAspLysProProArgGlyLysHisSerThrLysGlyLeuGlyLysThr 99
 1541 CCGAAGGATTGCTTCAAGGTAACATAGCACCAAGGGCTGGCAAGATG 1590
 100 ValProLeuGluSerGluPheValLysTyrArgAspValValValPr 116
 1591 GCTCCAGTCTGCGCCACTTCGTCACCTCAATGGAGT...ACAGTGCC 1637
 116 cCysGlyLysProValProSerSerIleArgSerSerGlu.....L 130
 1638 ATTAGGACCAGCAAGTGCACACAGGAATTCGAATCCAGATGGTTATACCC 1687
 130 eumetTyrAsnGluTyrIleValTyrAsnThrSerGlnValLysMetGln 146
 1688 TCAACTACAATGAATATATTGATATATAACCCCAACCAAGGTCGATGCGG 1737
 147 PheLeuLeuLysValArgPheHis 154
 1738 TACCTTTAAAGGTTCAAGTTTAAT 1761

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT: AAC85341

seq_documentation_block:

ID AAC85341 standard; cDNA; 3200 BP.

XX AAC85341;

AC AAC85341;

[illegible]

OM of: US-09-236-995D-4 to: Issued_Patents_NA:* out_format: pfs
Date: Mar 9, 2002 7:46 PM
About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=framet.p2n.model -DEV=xlp
-Q/cgn2_1/USPTO_spool/US09236995/runat_08032002_113149_25118/app_query.fasta.1.1273
-DB=Issued_Patents_NA -OFMT=fastp -SUFFIX=rn1 -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-GAPOP=4.000 -GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-XGAPOP=6.000 -XGAPEXT=7.000 -XGAPOP=10.000 -XGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=bloms62
-TRANS=human40 cdi -LIST=45 -DOALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM=ext -MINLEN=0 -MAXLEN=2000000000
-USER=US09236995_@CGN1_1_111 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPXY
-WAIT -THREADS=1

Search information block:

Query: US-09-236-995D-4
Query length: 157
Database: Issued_Patents_NA:*
Database sequences: 351203
Database length: 11323899
Search time (sec): 114.960000

score_list:

Sequence	Strd Orig	ZScore	EScore	Len
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-044-618-5	1 Documentation	503.50	1160.85	3747
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-044-618-7		435.50	993.59	5345
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-044-618-3		429.50	987.30	2682
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-09-328-111-216		358.50	834.70	595
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-044-618-4		350.50	803.96	1592
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-09-196-387-1		323.00	247.12	4134
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-09-196-387-7		323.00	246.15	4491
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-09-196-387-9		323.00	245.72	4657
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-522-229B-1		77.00	146.72	1776
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-09-102-433-1		75.00	146.72	1776
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-09-011-197-1		75.00	139.78	2361
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-09-371-696-1		72.50	135.91	1779
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-998-416-1056		71.50	147.56	538
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-298-426-3		69.50	128.74	1776
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-09-162-373-2		69.00	118.21	3930
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-09-467-946-2		69.00	118.21	3930
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-524-828-2		68.50	127.62	1593
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-975-114A-2		68.50	127.62	1593
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-849-281A-2		68.50	127.62	1593
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-524-828-1		68.50	123.57	2247
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-975-114A-1		68.50	123.57	2247
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-658-136-2		68.00	85.13	53526
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-658-136-1		68.00	85.12	53577
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-620-467A-6		67.00	136.11	569
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-348-572-6		67.00	136.11	569
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-09-041-090B-6		67.00	136.11	569
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-752-760A-1		67.00	87.69	33081
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-427-497E-4		66.00	115.86	83.71
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-427-497E-3		66.00	113.47	113.86
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-341-843B-1		66.00	111.49	146.76
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-427-497E-1		66.00	111.49	146.76
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-427-497E-2		66.00	111.49	146.76
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-09-232-200-36		65.50	113.44	114.19
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-09-232-200-56		65.50	113.44	114.19
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-09-232-197-36		65.50	113.44	114.19
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-09-232-197-56		65.50	113.44	114.19
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-453-695A-111		65.50	112.86	123.13
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-268-161A-111		65.50	112.86	123.13
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-453-702A-111		65.50	112.86	123.13
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-09-099-639-111		65.50	112.86	123.13

/cgn2_6/ptodata/2/ina/5A_COMB.seq:PCT-US95-08071-111 + 65.50 112.86 123.13
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-750-128-7 - 65.00 125.36 24.78
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-239-276-10 + 65.00 110.63 163.89
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-468-579B-10 + 65.00 110.63 163.89
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-468-577B-10 + 65.00 110.63 163.89
seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-044-618-5

seq_documentation_block:

; Sequence 5, Application US/08044618
; Patent No. 5449605
; GENERAL INFORMATION:
; APPLICANT: SMULSON, MARK
; TITLE OF INVENTION: METHOD OR DETECTING A PREDISPOSITION TO
; TITLE OF INVENTION: CANCER BY THE USED OF RESTRICTION FRAGMENT LENGTH
; TITLE OF INVENTION: POLYMORPHISM OF THE GENE FOR THE HUMAN POLY (ADP-RIBOSE)
; TITLE OF INVENTION: POLYMERASE
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1225 Connecticut Suite 300
; CITY: Washington
; STATE: D.C.
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/044.618
; FILING DATE: 19930406
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/257.696
; FILING DATE: 14-OCT-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: FOX, SAMUEL L
; REGISTRATION NUMBER: 30,353
; REFERENCE/DOCKET NUMBER: 0654.0490001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)466-0800
; TELEFAX: (202)833-8716
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3747 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; US-08-044-618-5

alignment_scores:

Quality: 503.50 Length: 156
Ratio: 3.702 Gaps: 1
Percent Similarity: 87.179 Percent Identity: 60.897

alignment_block:

US-09-236-995D-4 x US-08-044-618-5 ..
Align seq 1/1 to: US-08-044-618-5 from: 1 to: 3747
1 AsnlysmetleuThrpHisGlySerArgLeuThrAsnPhValGlyI1 17
2756 AACCGAGATGCTGTGGCAGGTCACGACCCACCACTTGTGGAT 2805
17 eLeuSerGlnGlyLeuArgIleAlaProProGluAlaProValThrGlyT 34
2806 CTTGTCCAGGGTCTCGGATAGCCGCGCTGAAGCCGCGTGCACAGCT 2855
34 yrMetPheGlyLysGlyLeuTyPheAlaAspLeuValSerLysSerAla 50
|||||


```
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/044,618
; APPLICATION NUMBER: US/08/044,618
; FILING DATE: 19930406
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/257,696
; FILING DATE: 14-OCT-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: FOX, SAMUEL L
; REGISTRATION NUMBER: 30,353
; REFERENCE/DOCKET NUMBER: 0654.0490001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)466-0800
; TELEFAX: (202)833-8716
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2682 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-044-618-3

alignment_scores:
  Quality: 429.50      Length: 156
  Ratio: 3.492        Gaps: 1
  Percent Similarity: 78.846  Percent Identity: 57.051

alignment_block:
US-09-236-995D-4 x US-08-044-618-3

Align seg 1/1 to: US-08-044-618-3 from: 1 to: 2682

1 AsnLysMetLeuLeuTrpHisGlySerArgLeuThrAsnPheValGlyI1 17
|||||
2050 AACTGAAGTTGCTGTGGCATGGTCCAGGACCACCACTTCTGGTGGAT 2099

17 eLeuSerGlnGlyLeuArgIleAlaProGluAlaProValThrGlyT 34
|||||
2100 CCTGTCCTGGTCTTGGATAGCCCTGCTCAAGCACCTGTGATGGGT 2149

34 yrMetPheGlyLysGlyLeuTyrrPheAlaAspLeuValSerLysSerAla 50
|||||
2150 ACATCTTTGGTAAAGTGATCTATTTTCGTGATCTTCTCCCAAGAGTGCC 2199

51 GlnTyrrCysTyrrValAspArgAsnAsnProValGlyLeuMetLeuSe 67
|||||
2200 ACGACTGCCATACATCTTAGGAAGACCAATAGGGTTANTCCTGTGCGA 2249

67 rGluValAlaLeuGlyAspMetTyrrGluLeuLysLysAlaThrSerMetA 84
|||||
2250 AGAAGTTGCCCTTGGAAGAGTGTGTGAAGTCAAGCATGCTTCACATATCA 2299

84 sPLysProArgGlyLysHisSerThrLysGlyLeuGlyLysThrVal 100
|||||
2300 GCAAGTTACCCAAAGGGCAAGCACAGTGTCAAAGTTTGGGCAAAACTACT 2349

101 ProLeuGluSerGluPheValLysTyrrPArgAspValValValProCy 117
|||||
2350 CTGACCTTTTACGTAGTATATCCCA...CTGGATGGTGTAGAGGTCCTCT 2396

117 sGlyLysProValProSerSerIleArgSerSerGluLeuMetTyrrAsnG 134
|||||
2397 TGGGACACGGGTTTCTATCTGGTGTGAATACACACTGTCTACTTATATG 2446

134 luTyrrIleValTyrrAsnThrSerGlnValLysMetGlnPheLeuLys 150
|||||
2447 AGTACATTGTCTATGATATATGCTCAGGTAATCTGAAATATCTGCTGAAA 2496
```

```
151 ValArgPheHisLys 156
|||||
2497 CTCAAATTCATTTAAG 2514

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-328-111-216

seq_documentation_block:
; Sequence 216, Application US/09328111
; Patent No. 6262333
; GENERAL INFORMATION:
; APPLICANT: Endege, Wilson O.
; APPLICANT: Steinmann, Kathleen E.
; APPLICANT: Astle, Jon H.
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Bushnell, Steven E.
; APPLICANT: Carroll III, Eddie
; APPLICANT: Catino, Theodore J.
; APPLICANT: Derti, Adnan
; APPLICANT: Ford, Donna M.
; APPLICANT: Lewis, Marcia E.
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; FILE REFERENCE: CCD-257 (US)
; CURRENT APPLICATION NUMBER: US/09/328,111
; CURRENT FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: US 60/088,801
; EARLIER FILING DATE: 1998-06-10
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 216
; LENGTH: 595
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(595)
; OTHER INFORMATION: n = A,T,C or G
US-09-328-111-216
```

```
alignment_scores:
  Quality: 358.50      Length: 108
  Ratio: 3.774        Gaps: 1
  Percent Similarity: 87.963  Percent Identity: 64.815

alignment_block:
US-09-236-995D-4 x US-09-328-111-216/rev

Align seg 1/1 to reverse of: US-09-328-111-216 from: 1 to: 595

1 AsnLysMetLeuLeuTrpHisGlySerArgLeuThrAsnPheValGlyI1 17
|||||
342 AACAGGATGCTTCTATGGCATGGTCCAGGATGAGTAAGTGGTGGGAAT 293

17 eLeuSerGlnGlyLeuArgIleAlaProGluAlaProValThrGlyT 34
|||||
292 CTGTAGCCATGGGCTTCGAATTCGCCACCTGAAGCTCCCATCAGGTT 243

34 yrMetPheGlyLysGlyLeuTyrrPheAlaAspLeuValSerLysSerAla 50
|||||
242 ACATGTTTGGGAAAGGAATCTACTTTGCTGACATGCTTCCAGAGTGCC 193

51 GlnTyrrCysTyrrValAspArgAsnAsnProValGlyLeuMetLeuSe 67
|||||
192 AATTACTGCTTTGCCCTCTCGCCTAAAGAAATACAGGACTGCTGCTTATC 143

67 rGluValAlaLeuGlyAspMetTyrrGluLeuLysLysAlaThr...SerM 83
|||||
142 AGAGGTAGCTAGGTAGTGTAGTGAATGAAGTACTAGAGGCAATCCTTAAG 93
```

83 eAspLysProProArgGlyLysHisSerThrLysGlyLeuGlyLysThr 99
::: |||||
92 CGAAGGATTGCTCAAGTAACAATAGCACCAAGGGGCTGGCAAGATG 43

100 ValProLeuGluSerGluPheVal 107
::: |||||
42 GTCGCCAGTTCTGCCACACTGCTC 19

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-044-618-4

seq_documentation_block:
; Sequence 4, Application US/08044618
; Patent No. 5449605
; GENERAL INFORMATION:
; APPLICANT: SMULSON, MARK
; TITLE OF INVENTION: METHOD OR DETECTING A PREDISPOSITION TO
; TITLE OF INVENTION: CANCER BY THE USE OF RESTRICTION FRAGMENT LENGTH
; TITLE OF INVENTION: POLYMORPHISM OF THE GENE FOR THE HUMAN POLY (ADP-RIBOSE)
; TITLE OF INVENTION: POLYMERASE
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1225 Connecticut Suite 300
; CITY: Washington
; STATE: D.C.
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/044,618
; FILING DATE: 19930406
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/257,696
; FILING DATE: 14-OCT-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: FOX, SAMUEL L.
; REGISTRATION NUMBER: 30,353
; REFERENCE/DOCKET NUMBER: 0654,0490001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)466-0800
; TELEFAX: (202)833-8716
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1592 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-044-618-4

alignment_scores:
Quality: 350.50 Length: 128
Ratio: 3.505 Gaps: 1
Percent Similarity: 78.125 Percent Identity: 58.594

alignment_block:

US-09-236-995D-4 x US-08-044-618-4

Align seg 1/1 to: US-08-044-618-4 from: 1 to: 1592

1 AsnLysMetLeuLeuTrpHisGlySerArgLeuThrAsnPheValGlyI 17
||| |||||
1209 AACTGAAGGTGCTGCTGGCATGGTCCAGACCACTTTGCTGGGAT 1258

17 eLeuSerGlnGlyLeuAlaProProGluAlaProValThrGlyT 34
||| |||||
1259 CCTGCTGCTGGGCTCTTTGGATAGACCCCTGCCTGAAGCACCTGTGATGGGCT 1308

34 yMetPheGlyLysGlyLeuTyrPheAlaAspLeuValSerLysSerAla 50
||| |||||
1309 ACATGTTTGTAAAGTGATCTATTTCGCTGATCTTCTCCTCAAGAGTGCC 1358

51 GlnTyrCysTyrValAspArgAsnProValGlyLeuMetLeuLeuSe 67
::: |||||
1359 AAGCACTGCCATACATCTTAGGAAGACCAATAGGTTAATCCTCTCGGA 1408

67 rGluValAlaLeuGlyAspMetTyrGluLeuLysLysAlaThrSerMet 84
||| |||||
1409 AGAAGTTGCCCTTGGAAACGTGTGGAAGTGAAGCATGCTTACATATCA 1458

84 sPlyProProArgGlyLysHisSerThrLysGlyLeuGlyLysThrVal 100
||| |||||
1459 GCAAGTTACCAAGGCAAGCAGTGTCAAAGTTTGGCAAACTACT 1508

101 ProLeuGluSerGluPheValLysTrpArgAspValValProcy 117
||| |||||
1509 CCTGACCTTTCAGCTAGTATCCCA...CTGGATGGTGTAGAGGTTCTCT 1555

117 sGlyLysProValProSerSerIleArgSerSer 128
||| |||||
1556 TGGGACCAGGTTTCATCTGTTGTAATGACACC 1589

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-196-387-1

seq_documentation_block:
; Sequence 1, Application US/09196387
; Patent No. 6277613
; GENERAL INFORMATION:
; APPLICANT: de Lange, Titia
; APPLICANT: Smith, Susan
; TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS
; TITLE OF INVENTION: OF USE THEREOF
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue, 4th Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/196,387
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/095,225
; FILING DATE: June 10, 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-230 CIP1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4134 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
US-09-196-387-1

[illegible]

APPLICANT: ROLOU, Lene Venke
APPLICANT: Andersen, Lene No. 5882911boe

Align seq 1/1 to: US-09-102-433-1 from: 1 to: 1776

```

48 sSer...AlaGlnTyrCysTyrValAspArgAsnAsnProVal.....61
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
291 CACCTTGAGCGAGTACATGGTGGTGCATTAATGGGACCAATCATTCACA 340
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
62 .GlyLeuMetLeuLeuSerGluValAlaLeuGlyAspMetTyrGluLeu 77
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
341 TGGCGACATATATCACTCGCGAGCGGTCAATCGGAGCTGGCGTTCATC 390
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
78 LysLysAlaThrSerMetAspLysProProArgGlyLysHisSerThr...93
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
391 GCTCGACTGAATTCGACCTGCTACCGAGGAGCGCTTGGCGACGTT 440
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
94 .....LysGlyLeuGlyLysThrValProLeuGluSerGluP 106
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
441 TCACCACCGCTGACGGAGCTG.....CCATTGAGGATCAGA 478
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
106 heValLysTrpArgaspValValProCysGlyLysProValPro 122
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
479 TGTGTTTTTGG.....TCGGCAGTGAACCCCGCAGCA 510
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
123 SerSerIleArgSerSerGluLeu 130
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
511 AGTTCTACTAGAGCGAGCTTTA 534
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-011-197-1

```

seq_documentation_block:
; Sequence 1, Application US/09011197
; Patent No. 6171789
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: No. 6171789el Insertion Sequence from a Virulent
; TITLE OF INVENTION: Isolate of Burkholderia Cepacia, and Diagnostic and
; TITLE OF INVENTION: Identification Procedures Based Thereon.
; NUMBER OF SEQUENCES: 16
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/002,398
; FILING DATE: 17-AUG-1995
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2361 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Burkholderia cepacia
; INDIVIDUAL ISOLATE: ET12/cbla
US-09-011-197-1

```

```

alignment_scores:
  Quality: 75.50      Length: 180
  Ratio: 0.910       Gaps: 10
  Percent Similarity: 46.111  Percent Identity: 23.889

```

alignment_block:
US-09-236-995D-4 x US-09-011-197-1 ..

Align seg 1/1 to: US-09-011-197-1 from: 1 to: 2361

```

1 AsnLysMetLeuLeuTrpHisGlySerArg.....Le 11
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1404 AACCTCGTGGCCAAAGTGGACCGGCTCTCGGCACGATTGGAGAGCGGAT 1453
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
11 uThrAsnPhe.....ValGlyIle.....17
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1454 GACCCAGTTTCGGCTGCTTTACCCCGAACGATTCAACATTGGAATCTGAA 1503
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
18 .....LeuSerGlnGlyLeu ArgIleAlaProProGlu.....28
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1504 TCTCAACCGGCTCACACAGGAATTCGGATACCTCCACCCACGACGCGC 1553
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
29 .....AlaProValT 32
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1554 GCGCGAGAAAAACCCAGCGCGCTGCTGTTTCGAATCGCGCGCGGTGA 1603
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
32 hrGlyTyrMetPhe.....GlyLysGlyLeuTyrPheAlaAspLeuVal 46
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1604 CCGCATCTCTGTTCTCAAGACCGGCTGCGCTGGCGGACCTGCC 1653
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
47 SerLys.....SerAlaGlnTyrCysTyr.....54
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1654 GCCGAGATGGGCTGCGGCTCGGGCGTGACTTGCTGGCGGACGGCTACGCGA 1703
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
55 .....ValAspArgAsnAsnProValGlyLeuMetL 65
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1704 TTGGCAGGCTGCGGGCGTATGGGATCGCCTGCACGAA.....TTGC 1744
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
65 euLeuSerGluValAlaLeuGlyAspMetTyrGluLeuLysLysAlaThr 81
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1745 TGCTTGGGAAGCTCGGAGCAGCAGACCAAAATCGATTTCTCGCGAGCGCA 1794
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
82 SerMetAspLysProProArgGlyLysHisSerThrLysGlyLeuGly 98
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1795 GTCGATTCTCTCATCGATTCGCGCGCTGGGGCGGGCCCAAAACTGGGCCA 1844
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
98 sThrValProLeuGluSerGluPheValLysTrpArgAspAspValVal 115
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1845 AACCCACCGCATCGCG.....CGGACGAGGTTCTAA 1876
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
115 alProCysGlyLysProValProSerSerIleArgSer 127
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1877 GCACCATCGTCCAGCAGCGCAATGCGCGCGCTCG 1914
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-371-696-1

```

seq_documentation_block:
; Sequence 1, Application US/09371696
; Patent No. 6287777
; GENERAL INFORMATION:
; APPLICANT: Sytkowski, Arthur J.
; APPLICANT: Yang, Meiheng
; TITLE OF INVENTION: NOVEL NPG-1 GENE THAT IS DIFFERENTIALLY EXPRESSED IN PROSTATE
; TITLE OF INVENTION: TUMORS
; FILE REFERENCE: 01948/053002
; CURRENT APPLICATION NUMBER: US/09/371,696
; CURRENT FILING DATE: 1999-08-10
; EARLIER APPLICATION NUMBER: US 09/022,238
; EARLIER FILING DATE: 1998-02-11
; EARLIER APPLICATION NUMBER: US 08/644,326
; EARLIER FILING DATE: 1996-05-10
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1779
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (236)...(1225)
; NAME/KEY: misc_feature
; LOCATION: (1)...(1779)

```


124 r.....lleargSerSerGluLeuMetTyrAsnGluTyrI 136
442 CTAGGCGCGGGCGCGCGGACCATTTAGTTCGACNGAATCNCATG 491
136 le 136
492 TC 493

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-298-426-3

seq_documentation_block:
: Sequence 3, Application US/08298426
: Patent No. 5516674
: GENERAL INFORMATION:
: APPLICANT: Roe, Richard M.
: APPLICANT: Hodgson, Ernest
: APPLICANT: Rose, Randy L.
: TITLE OF INVENTION: Methods and Compositions for the
: Degradation of Insecticides and Chlorinated Organics
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Kenneth D. Sibley; Bell, Seltzer, Park &
: ADDRESS: Gibson
: STREET: Post Office Drawer 34009
: CITY: Charlotte
: STATE: No. 5516674th Carolina
: COUNTRY: USA
: ZIP: 28234
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/298,426
: FILING DATE:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Sibley, Kenneth D.
: REGISTRATION NUMBER: 31,665
: REFERENCE/DOCKET NUMBER: 5051-207A
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 919-420-2200
: TELEFAX: 919-881-3175
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1776 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: ORIGINAL SOURCE:
: ORGANISM: Heliothis virescens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 103..1680
: US-08-298-426-3

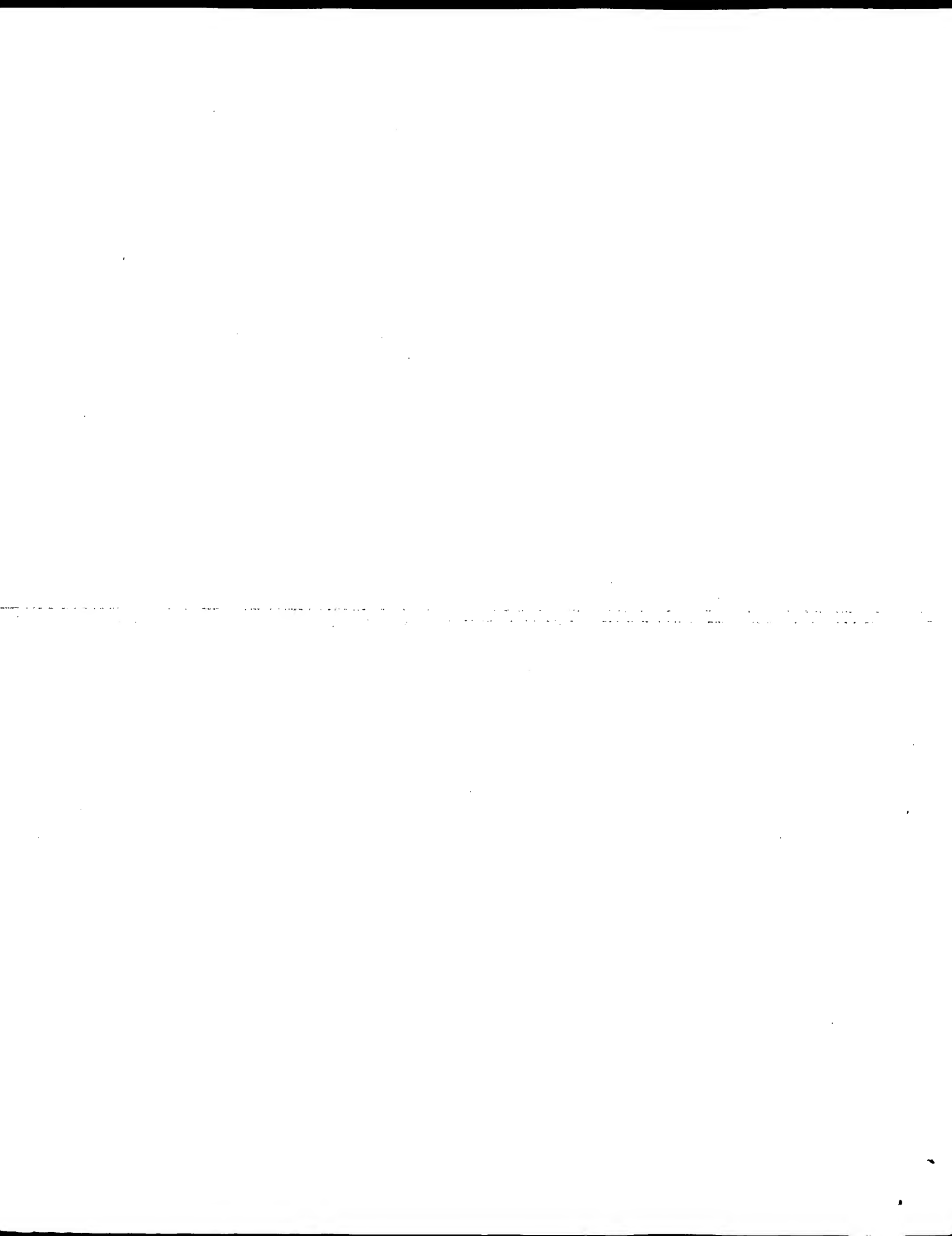
alignment_scores:
Quality: 69.50 Length: 202
Ratio: 0.808 Gaps: 11
Percent Similarity: 42.574 Percent Identity: 22.277

alignment_block:
US-09-236-995D-4 x US-08-298-426-3 ..
Align seg 1/1 to: US-08-298-426-3 from: 1 to: 1776
4 LeuLeuTrpHisGlySerArgLeuThrAsnPheValGlyIleLeuSerGl 20
|||||
211 CTGATGGTGGTGGTATGACGAGGACACTTCGTTGATACCCCTGGGGCG 260

20 nGlyLeuArgIleAlaProGluAlaProValThrGlyTyrMetPheG 37
: ||| :
261 GTGTGTCAAGGCTTTCTCTGATGATAAGTAGTAGGACACTAC 303
: ||| :
37 lyGlyLeuTyrPheAlaAspLeuValSer..... 47
: ||| :
304GACATGGTGGAGCCCTATCTTGGTAGTTCTC 333
: ||| :
48LysSerAlaGlnTyrCysTy 54
: ||| :
334 GACGTGGGACACTGTCAAGAGGATCACCCTCAAGACATTTGAACAT...TT 380
: ||| :
54 rValAspArgAsn.....AsnProVal.....G 62
: ||| :
381 CGTTGACAGGGAACCTTCACCAGAGCTTTGATCCCATCTTTGGAAAGAG 430
: ||| :
62 lyLeuMetLeuLeuSerGluValAlaLeuGlyAspMetTyrGluLeuLys 78
: ||| :
431 GGCTGCTCTTGTCTACAT.....GCTGACGAATGGAAGCAATG 468
: ||| :
79 LysAlaThr..... 81
: ||| :
469 CGGTCTAGCATGAGTCCAGCGCTTCACGAGCTCCAAAGATGCGCCTGATGGT 518
: ||| :
82SerMetAspLysProProArgGlyL 90
: ||| :
519 GCCCTTCATGGAGAGATCGCTTTGGAAATGATTAGTAGTACTCCGGGGGA 568
: ||| :
90 ysHisSerThrLysGlyLeuGlyLys.....ThrValProLeuGlu 103
: ||| :
569 AG.....ATCAAGGATTTCTGGGAAACCTTACATCGAGCTGGAAGCCAAG 612
: ||| :
104 SerGluPheValLysTrpArgAspValValProCys.....G 118
: ||| :
613 AGTATGATGACCGAGTACGCGAATGAGCTATAGCCTCATCGCCTCTCGG 662
: ||| :
118 yLysProValProSerSerIleArgSerSerGluLeuMetTyrAsnGlu 135
: ||| :
663 GTTGAAGTGAACCTCCAGCGGCTCGGACACGAGTTTATGTCAACAGTC 712
: ||| :
135 yrlleValTyrAsnThrSerGlnValLysMetGlnPheLeuLeuLysVal 151
: ||| :
713 AAGCTATC.....ACCAAGTTTAAGTTTTCAGCCTTTCTGAAAGTC 753
: ||| :
152 ArgPhe 153
: ||| :
754 CTGTTC 759

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-09-162-373-2

seq_documentation_block:
: Sequence 2, Application US/09162373
: Patent No. 6013454
: GENERAL INFORMATION:
: APPLICANT: Tang, Y. Tom
: APPLICANT: Corley, Neil C.
: APPLICANT: Guegler, Karl J.
: APPLICANT: Patterson, Chandra
: TITLE OF INVENTION: KINESIN-LIKE MOTOR PROTEIN
: FILE REFERENCE: PF-0593 US
: CURRENT APPLICATION NUMBER: US/09/162,373
: CURRENT FILING DATE: 1998-09-28
: NUMBER OF SEQ ID NOS: 2
: SOFTWARE: PERL Program
: SEQ ID NO 2
: LENGTH: 3930
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: 1281811
: US-09-162-373-2



OM of: US-09-236-995d-4 to: EST:* out_format : pfs

Date: Mar 9, 2002 8:22 PM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

```
-MODEL=framer_p2n.model -DEV=xlp
-O=/cqn2.1/USPTO_spoil/US09236995/runat_08032002.113148.25093/app_query.fasta.1.1273
-DB=EST -OEXT=fastap -SUFFIX=rst -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPEXT=0.000 -LOOPEXT=0.000 -GAPOP=4.500
-QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-DEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-DEXT=7.000 -START=1 -MATRIX=blousum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0
-MAXLEN=200000000 -USER=US09236995_@CGN1.1_5291 -NCPU=6
-ICPU=3 -LONGLOG -NO_XLUPY -WAIT -THREADS=1
```

Search information block:

Query: US-09-236-995d-4

Query length: 157

Database: EST:*

Database sequences: 11351937

Database length: 1077921985

Search time (sec): 2336.950000

score_list:

Sequence	Strd	Orig	Zscore	EScore	Len	Documentation
gb_est2:BF255013	+	714.00	1548.29	4.5e-77	607	BF255013 HVSM012 AL2R000101 IREU
gb_est1:BE419432	+	660.00	1431.16	1.5e-70	536	BE419432 HVSM012 AL2R000101 IREU
gb_est1:AV834168	-	642.00	1389.55	3.1e-68	661	AV834168 AV834168 K. Sato unpub
gb_est1:BE420229	+	569.00	1235.46	1.2e-59	369	BE420229 WMS04 C11R000101 IREU
gb_est2:BF217470	+	518.50	1117.95	4.2e-53	716	BF217470 602933675F1 NCI_CGAP_1
gb_est1:BE304535	+	517.50	1118.08	4.1e-53	571	BE304535 601105502F1 NIH_MGC_1
gb_est1:AA397988	+	517.50	1117.93	4.2e-53	579	AA397988 zt86c12.r1 Soares_test
gb_est1:AA401836	+	517.50	1116.46	5.0e-53	669	AA401836 zt86c12.r1 Soares_test
gb_est2:BF094336	+	517.50	1113.24	7.6e-53	916	BF094336 602859629F1 NIH_MGC_1
gb_est2:BF793705	+	517.50	1112.85	8.0e-53	952	BF793705 602255121F1 NIH_MGC_1
gb_est2:BF691127	+	513.50	1109.02	1.3e-52	587	BF691127 601662539F1 NCI_CGAP_1
gb_est1:BE382739	+	513.50	1100.51	3.9e-52	709	BE382739 601297572F1 NIH_MGC_1
gb_est2:BF079902	+	509.50	1098.09	5.3e-52	725	BF079902 602876096F1 NCI_CGAP_1
gb_est2:BF675913	+	503.50	1082.41	4.0e-51	928	BF675913 603622263F1 NCI_CGAP_1
gb_est2:BF6042229	+	493.00	1065.68	3.4e-50	502	BF6042229 su3c12.y1 Gm-cl055 GI
gb_est2:BF6280821	+	489.50	1052.48	1.8e-49	861	BF6280821 602401166F1 NIH_MGC_1
gb_est1:BE729084	+	488.50	1052.59	1.8e-49	688	BE729084 601561120F1 NIH_MGC_1
gb_est2:BF031594	+	485.50	1042.28	6.8e-49	990	BF031594 602299739F1 NIH_MGC_1
gb_est2:BF026745	+	479.50	1029.29	3.6e-48	975	BF026745 601671953F1 NIH_MGC_1
gb_est2:BF101770	+	473.50	1019.85	1.2e-47	678	BF101770 601753112F1 NCI_CGAP_1
gb_est2:BF6941860	+	470.00	1017.10	1.7e-47	419	BF6941860 EST421439 MGHG Medicag
gb_est1:AV833893	-	469.50	1011.34	3.6e-47	661	AV833893 AV833893 K. Sato unpub
gb_est2:BF0429402	+	465.50	1003.33	1.1e-46	632	BF0429402 602499622F1 NIH_MGC_1
gb_est2:BF042005	+	455.50	982.34	1.5e-45	560	BF042005 BP250013A20B10 Soares
gb_est2:BF769270	-	450.50	974.20	4.2e-45	425	BF769270 OV4-IT0004-111100-473
gb_est1:AW144795	-	447.50	966.23	1.2e-44	487	AW144795 EST291865 Normalized
gb_est2:BF66030	+	445.50	959.83	2.7e-44	593	BF66030 zc56e05.r1 Soares_parrat
gb_est2:BF0303304	-	445.50	957.53	3.6e-44	742	BF0303304 f151b10.x1 Sugano Kawa
gb_est2:BF0393399	-	445.50	957.25	3.7e-44	763	BF0393399 602411745F1 NIH_MGC_1
gb_est2:BF0287484	+	443.00	948.26	1.2e-43	1075	BF0287484 602382082F1 NIH_MGC_1
gb_est1:AA45188	-	442.00	950.05	9.4e-44	728	AA45188 LD21673.3prime LD broa
gb_est1:AV708160	+	441.50	950.04	9.4e-44	655	AV708160 AD708160 ADC Homo sapi
gb_est2:BF751755	+	441.00	948.37	1.2e-43	771	BF751755 602730481F1 NIH_MGC_1
gb_est1:AA263755	-	441.00	950.49	8.9e-44	563	AA263755 LD07027.5prime LD Dro
gb_est1:AA1417842	-	435.50	937.73	4.6e-43	603	AA1417842 tg72a07.x1 Soares_Nhm
gb_est1:BF564253	+	434.50	933.55	7.8e-43	732	BF564253 601343054F1 NIH_MGC_1
gb_est2:BF884815	+	433.50	929.28	1.3e-42	897	BF884815 601510541F1 NIH_MGC_1
gb_est2:BF793250	+	429.50	920.35	4.2e-42	911	BF793250 602252914F1 NIH_MGC_1
gb_est2:BF119670	+	428.50	921.87	3.5e-42	634	BF119670 602673286F1 NIH_MGC_1
gb_est1:AA212857	+	422.00	907.90	2.1e-41	617	AA212857 mw84d01.r1 Soares_mous
gb_est1:BE270258	+	416.00	894.30	1.2e-40	644	BE270258 601186052F1 NIH_MGC_1

gb_est1:BE488132 + 414.00 892.53 1.5e-40 499 ! BE488132 177421 BARC 5BOV Bo
gb_est1:AI202751 - 411.50 885.44 3.7e-40 584 ! AI202751 q145c05.x1 NCI_CGAP
gb_est1:AA129412 + 409.00 881.71 6.0e-40 492 ! AA129412 zn83c06.r1 Stratage
gb_est2:BG476961 + 407.50 875.00 1.4e-39 687 ! BG476961 602522176F1 NIH_MGC

seq_name: gb_est2:BF255013

seq_documentation_block:

LOCUS BF255013 607 bp mRNA EST 23-FEB-2001
DEFINITION HVSM012005L18f Hordeum vulgare seedling root EST library HVCDNA0007
(etiolated and unstressed) Hordeum vulgare cDNA clone
HVSM012005L18f, mRNA sequence.

ACCESSION BF255013

VERSION BF255013

KEYWORDS EST.

SOURCE barley.

ORGANISM Hordeum vulgare

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
; Triticeae; Hordeum.

REFERENCE 1 (bases 1 to 607)

AUTHORS Wing, R., Close, T.J., Kleinhofs, A., Wise, R., Begum, D., Frisch, D., Yu

, Y., Anderson, H., Dale, J., Henry, D., Kernodle, S., Palmer, M., Rambo
, T., Saski, C., Schwartzbeck, J., Simmons, J., Choi, D.W., Main, D. and
Wood, T.

Development of a genetically and physically anchored EST resource

for barley genomics

Unpublished (2000)

On Nov 16, 2000 this sequence version replaced gi:11184130.

JOURNAL Contact: Wing RA

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Tel: 864 656 7288

Fax: 864 656 4293

Email: rwing@clemson.edu

Seq primer: AATTAACCTCACTAAGGG

High quality sequence stop: 551.

FEATURES

Location/Qualifiers

1..607

/organism="Hordeum vulgare"

/cultivar="Morex"

/db_xref="taxon:4513"

/clone="HVSMEF0005L18f"

/clone.lib="Hordeum vulgare seedling root EST library

HVCNA0007 (etiolated and unstressed)"

/tissue_type="Seedling root"

/lab_host="TJC121"

/note="Vector: lambdaZAP; Site_1: EcoRI; Site_2: XhoI; For
more details on library preparation and sequence analysis
see http://www.genome.clemson.edu/projects/barley/ To
order a clone see http://www.genome.clemson.edu/orders"

BASE COUNT 165 a 113 c 163 g 166 t

ORIGIN

alignment_scores:

Quality: 714.00 Length: 146

Ratio: 4.993 Gaps: 0

Percent Similarity: 97.945 Percent Identity: 91.781

alignment_block:

US-09-236-995d-4 x BF255013

Align seg 1/1 to: BF255013 from: 1 to: 607

12 ThrAsnPhaValGlyIleLeuSerGlnGlyLeuArgIleAlaProGln 28

1 ACAAATTTATTGGAATCTTATGTAAGGACATAGGACCTCTCTGA 50

28 uAlaProValThrGlyTyrMetPheGlyLysGlyLeuTyrPheAlaAspL 45

|||||

51 GGACCCCGTAGGGGCTATATGTTTGGCAAGGCTCTTCTTTCAGATT 100

45 euValSerLysSerAlaGlnTyrCysTyrValAspArgAsnProVal 61
 101 TAGTAAGCAAGAGTCACAGTATTCTTATGTGGATAGAAATAATCCGAGT 150
 62 GlyLeuMetLeuSerGluValAlaLeuGlyAspMetTyrGluLeuLeu 78
 151 GCCTTGATGCTCTTCTTCAGGTTCTCTAGGAGACATGCATGAAGTAA 200
 78 slsAlaThrSerMetAspLysProProArgGlyLysHisSerThrLysG 95
 201 AAAAGCAACGCCAATGGCAACACCTCCAAAGAGGAAGCATTCGACCAAGG 250
 95 LyLeuGlyLysThrValProLeuGluSerGluPheValLysTrpArgAsp 111
 251 GCTTAGGCAAACTGTGGCACTAGAGTCGGAGTTTGTAAATGGAGGAT 300
 112 AspValValProCysGlyLysProValProSerSerIleArgSerSe 128
 301 GATGTCGTCGTCCTGTGGCAAGCAGTCCAGCATCATCAGGGCATC 350
 128 rGluLeuMetTyrAsnGluTyrIleValTyrAsnThrSerGlnValLysM 145
 351 TGAGCTTCTGTACAACGAGTATATAGTACAAACACAGCTCAGGTGAAGA 400
 145 etGlnPheLeuLeuLysValArgPheHisHisLysArg 157
 401 TGCAGTCTTGTGAAGGTCAAAATTCCTCACAAGCGT 438

seq_name: gb_est1:BE419432

seq_documentation_block: 536 bp mRNA EST 24-JUL-2000
 LOCUS BE419432
 DEFINITION WWS012.A12R000101 ITEC WWS Wheat Scutellum Library Triticum aestivum cDNA clone WWS012.A12, mRNA sequence.

ACCESSION BE419432

VERSION BE419432.1 GI:9417278

KEYWORDS EST.

SOURCE bread wheat.

ORGANISM Triticum aestivum

REFERENCE 1 (bases 1 to 536)
 AUTHORS Anderson, O.A., Appels, R., Bailey, P., Blake, T., Close, T., Cloutier, S., Dubcovsky, J., Feuillet, C., Gale, M., Graner, A., Gustafson, P., Herrmann, R.G., Holton, T., Jacquemin, J.M., Jia, J., Joudrier, P., Langridge, P., Lazo, G.R., Lin, J.J., McGuire, P., Ogihara, Y., Pecchioni, N., Qualset, C., Schuch, W., Selvaraj, G., Shariflou, M., Sorrells, M., Warburton, M. and Wenzel, G.
 TITLE International Triticaceae EST Cooperative (ITEC): Production of Expressed Sequence Tags for Species of the Triticaceae
 JOURNAL Unpublished (2000)
 COMMENT Contact: Schuch W
 Zeneca Wheat Improvement Centre, Norwich Research Park
 Colney Lane, Norwich NR4 7UH UNITED KINGDOM
 Tel: 44 1603 250 2600
 Fax: 44 1603 250 699
 Email: wolfgang.schuch@aguk.zeneca.com
 International Triticaceae EST Cooperative (ITEC)
 http://wheat.pw.usda.gov/genome.

FEATURES

source 1..536
 /organism="Triticum aestivum"
 /cultivar="Novosibirskaya 67"
 /db_xref="taxon:4565"
 /clone="WWS012.A12"
 /clone_lib="ITEC WWS Wheat Scutellum Library"
 /tissue_type="scutellum callus"
 /note="M13 Reverse sequencing primer used for 5' end of clone."

BASE COUNT 143 a 106 c 144 g 139 t 4 others

ORIGIN

alignment_scores:
 Quality: 660.00 Length: 134
 Ratio: 5.000 Gaps: 0
 Percent Similarity: 98.507 Percent Identity: 91.791

alignment_block:

US-09-236-995D-4 x BE419432 ..

Align seg 1/1 to: BE419432 from: 1 to: 536

24 ileAlaProProGluAlaProValThrGlyTyrMetPheGlyLysGlyLe 40
 11 ATAGCACCTCCTCAGGCTCCTGTGACAGGCTATATGTTGGCAAGGCCT 60
 40 uTyrPheAlaAspLeuValSerLysSerAlaGlnTyrCysTyrValAsp 57
 61 CTACTTTGCAGATTAGTAAGCAAGAGTGCACAGTATTGTTATGTGATA 110
 57 rGAsnAsnProValGlyLeuMetLeuLeuSerGluValAlaLeuGlyAsp 73
 111 GAAATAATCCGACTGGCTTCATGCTTCTTCTGAGGTTGCTCTAGGAGAC 160
 74 MetTyrGluLeuLysLysAlaThrSerMetAspLysProProArgGlyLy 90
 161 ATGCATGAACCTGAAAAGGCCAGCCCAATGACAAACCTCCAAAGAGGAAA 210
 90 sHisSerThrLysGlyLeuGlyLysThrValProLeuGluSerGluPheV 107
 211 GCATTCGACCAAGGCTTAGGAAAACTGTGCGCTAGAGTCGGAGTTTG 260
 107 allyTyrTrpArgAspAspValValProCysGlyLysProValProSer 123
 261 TGAATGGAGGATGATATCTGCTGCTTGTGGCAAGCCAGTCCAGCA 310
 124 SerIleArgSerSerGluLeuMetTyrAsnGlnTyrIleValTyrAsnTh 140
 311 TCATATCAGGCATCTGAGCTTCTGTACACGAGTATATAGTCTACAACAC 360
 140 rSerGlnValLysMetGlnPheLeuLeuLysValArgPheHisHisLys 157
 361 AGCCAGGTGAAGATCGATTCTTGTGAAGGTGAGTCCGTCACCAAGC 410
 157 rg 157
 411 GT 412

seq_name: gb_est1:AV834168

seq_documentation_block:

LOCUS AV834168 661 bp mRNA EST 22-JUN-2001
 DEFINITION AV834168 K. Sato unpublished cDNA library: Hordeum vulgare subsp. vulgare shoots germination Hordeum vulgare subsp. vulgare cDNA clone rbags5k04, mRNA sequence.

ACCESSION AV834168

VERSION AV834168.1 GI:14526257

KEYWORDS EST.

SOURCE Hordeum vulgare subsp. vulgare.

ORGANISM Hordeum vulgare subsp. vulgare.

REFERENCE 1 (bases 1 to 661)
 AUTHORS Sato, K.
 TITLE Barley EST sequencing project in NIG and Okayama Univ
 JOURNAL Unpublished (2001)
 COMMENT Contact: Kazuhiro Sato
 Research Institute for Bioresources
 Okayama University, Barley Germplasm Center
 Chuo 2-20-1, Kurashiki, Okayama 710-0046, Japan
 Email: kzsato@rib.okayama-u.ac.jp,
 URL: http://www.rib.okayama-u.ac.jp/barley/

Sato, K., Saisho, D., Takeda, K., Shini, T. and Kohara, Y. Direct submission;
database: <http://www.shigen.nig.ac.jp/barley/Barley.html>.

FEATURES

Location/Qualifiers
1. 661
/organism="Hordeum vulgare subsp. vulgare"
/cultivar="Haruna Nijo"
/db_xref="taxon:112509"
/clone_lib="K. Sato unpublished cDNA library: Hordeum vulgare subsp. vulgare shoots germination"
/tissue_type="shoots"
/dev_stage="germination"
BASE COUNT 180 a 181 c 132 g 168 t
ORIGIN

alignment_scores:

Quality: 642.00 Length: 136
Ratio: 4.827 Gaps: 1
Percent Similarity: 97.794 Percent Identity: 90.441

alignment_block:

US-09-236-995D-4 x AV834168/rev ..

Align seg 1/1 to reverse of: AV834168 from: 1 to: 661

23 ArgileAlaProGluAlaProValThr.GlyTyrMetPheGlyLysG 39
|||||
661 AGGATAGCACCCTCTGAGGACCCGCTGACCGGGCTATGTTGGCAAAG 612
39 lYLeuTyrPheAlaAspLeuValSerLysSerAlaGlnTyrCysTyrVal 55
|||||
611 GCCTCTACTTTGACATTTAGTAGCAGAGTGCACAGTATTGTTATG 562
56 AspArgAsnProValGlyLeuMetLeuLeuSerGluValAlaLeuG 72
|||||
561 GATAGAAAAATCCGACTGGCTTCTCTCTGAGGTTGCTCTAGG 512
72 yAspMetTyrGluLeuLysLysAlaThrSerMetAspLysProProArg 89
|||||
511 AGACATTTCATGAACCTGAAAAAGCAACGCCAATGGACAAACCTCCAAG 462
89 lYLeuHisSerThrLysGlyLeuGlyLysThrValProLeuGluSerGlu 105
|||||
461 GGAAGCATTCGACCAAGGCTTAGGCAAACTGTGCCACTAGAGTCGAG 412
106 PheValLysTrpArgAspValValProCysGlyLysProValPr 122
|||||
411 TTTGTTAAATGGAGGATGATGCTGCTGCTTGTGGCAAGCCAGTGCC 362
122 oSerSerIleArgSerSerGluLeuMetTyrAsnGluTyrIleValTyrA 139
|||||
361 AGCATCTATCAGGCGCTCTGAGCTTCTGTACACAGGATATATAGTGTA 312
139 snThrSerGlnValLysMetGlnPheLeuLeuLysValArgPheHisHis 155
|||||
311 ACACAGCTCAGGTGAGATGCGATGCTTCTGTGAAGTCAAAATCCGTCAC 262
156 LysArg 157
|||||
261 AAGCGT 256

seq_name: gb_estl:BE420229

seq_documentation_block:

LOCUS BE420229 369 bp mRNA EST 24-JUL-2000
DEFINITION WWS04.C11R000101 ITEC WWS Wheat Scutellum Library Triticum aestivum
CDNA clone WWS04.C11, mRNA sequence.
ACCESSION BE420229
VERSION BE420229.1 GI:9418075
KEYWORDS EST.
SOURCE bread wheat.

ORGANISM

Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Triticum.

REFERENCE
AUTHORS

1 (bases 1 to 369)
Anderson, O.A., Appels, R., Bailey, P., Blake, T., Close, T., Cloutier,
S., Dubcovsky, J., Feuillet, C., Gale, M., Graner, A., Gustafson, P.,
Herrmann, R.G., Holton, T., Jacquemin, J.M., Jia, J., Joudrier, P.,
Langridge, P., Lazo, G.R., Lin, J.J., McGuire, P., Ogihara, Y.,
Pecchioni, N., Qualset, C., Schuch, W., Selvaraj, G., Shariflou, M.,
Sorrells, M., Warburton, M. and Wenzel, G.

TITLE

International Triticeae EST Cooperative (ITEC): Production of
Expressed Sequence Tags for Species of the Triticeae

JOURNAL
COMMENT

Unpublished (2000)
Contact: Schuch W
Zeneca Wheat Improvement Centre, Norwich Research Park
Colney Lane, Norwich NR4 7UH UNITED KINGDOM
Tel: 44 1603 250 2600
Fax: 44 1603 250 699
Email: wolfgang.schuch@aguk.zeneca.com
International Triticeae EST Cooperative (ITEC)
<http://wheat.pw.usda.gov/genome>.

FEATURES

source

1. 369
/organism="Triticum aestivum"
/cultivar="Novosibirskaya 67"
/db_xref="taxon:4565"
/clone="WWS04.C11"
/clone_lib="ITEC WWS Wheat Scutellum Library"
/tissue_type="scutellum callus"
/note="M13 Reverse sequencing primer used for 5' end of
clone."

BASE COUNT 103 a 78 c 101 g 87 t
ORIGIN

alignment_scores:

Quality: 569.00 Length: 120
Ratio: 4.822 Gaps: 0
Percent Similarity: 98.333 Percent Identity: 91.667

alignment_block:

US-09-236-995D-4 x BE420229 ..

Align seg 1/1 to: BE420229 from: 1 to: 369

24 IleAlaProGluAlaProValThrGlyTyrMetPheGlyLysGlyLe 40
|||||
11 ATAGCACCTCCTGAGGCTCCTGTGACAGCTATATGTTGGCAAGGCT 60
40 uTyrPheAlaAspLeuValSerLysSerAlaGlnTyrCysTyrValAsp 57
|||||
61 CTACTTGCAGATTTAGTAAGCAAGAGTGCACAGTATTGTTATGTTGATA 110
57 rgAsnAsnProValGlyLeuMetLeuLeuSerGluValAlaLeuGlyAsp 73
|||||
111 GAAAAATCCGACTGCTGATGCTTCTTCTGAGGTTGCTCTAGGAGAC 160
74 MetTyrGluLeuLysLysAlaThrSerMetAspLysProProArgGly 90
|||||
161 ATGCATGAACCTGAAAAAGGCCACGCCAATGGACAAACCTCCCAAGGAAA 210
90 sHisSerThrLysGlyLeuGlyLysThrValProLeuGluSerGluPheV 107
|||||
211 GCATTCACCAAGGCTTAGGAAAACTGTCCGCTAGAGTCGGAGTTTG 260
107 alLysTrpArgAspValValProCysGlyLysProValProSer 123
|||||
261 TGAATGAGGAGTATGATATCGTCGCTTGTGGCAAGCCAGTGCAGCA 310
124 SerIleArgSerGluLeuMetTyrAsnGluTyrIleValTyrAsnTh 140
|||||
311 TCTATCAGGCGCATCTGAGCTTCTGTC.AACGAGTATATAGTACAAAC 359

140 rSerGlnVal 143
 1:::|||||
 360 AGCCAGCGTG 369

seq_name: gb_est2:B1217470

seq_documentation_block:

LOCUS B1217470 716 bp mRNA EST 11-JUL-2001
 DEFINITION 602933675F1 NCI_CGAP_Li9 Mus musculus cDNA clone IMAGE:5096850 5',
 mRNA sequence.

ACCESSION B1217470
 VERSION B1217470.1 GI:14670914

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 1 (bases 1 to 716) Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

AUTHORS NIH-MGC http://mgc.nci.nih.gov/.

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM11232 row: f column: 19

High quality sequence stop: 696.

FEATURES

source

1..716

/organism="Mus musculus"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone="IMAGE:5096850"

/clone_lib="NCI_CGAP_Li9"

/lab_host="DH10B (T1 phage-resistant)"

/note="Organ: liver; Vector: pCMV-SPORT6; Site_1: NotI;

Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.

Average insert size 1.9 kb. Constructed by Life

Technologies. Note: this is a NCI_CGAP Library."

175 a 196 c 191 g 154 t

alignment_scores:

Quality: 518.50 Length: 156

Ratio: 3.812 Gaps: 1

Percent Similarity: 87.179 Percent Identity: 62.179

alignment_block:

US-09-236-995d-4 x B1217470

Align seg 1/1 to: B1217470 from: 1 to: 716

1 AsnLysMetLeuLeuTrpHisGlySerArgLeuThrAsnPhValGlyI 17

||||| ||||||| ||||||| ||||||| ||||||| |||||||

12 AACCGGAGGCTCTGTGGCAGCGCTCCAGGACCACCAACTTTGTGGCAT 61

||||| ||||||| ||||||| ||||||| ||||||| |||||||

17 eLeuSerGlnGlyLeuArgIleAlaProProGluAlaProValThrGlyT 34

||||| ||||||| ||||||| ||||||| ||||||| |||||||

62 CTTGTGCGAGGCTGTGCGGATAGCCCCACCTGAAGCGCTGTGACAGGCT 111

||||| ||||||| ||||||| ||||||| ||||||| |||||||

34 yrMetPheGlyLysGlyLeuTyrPheAlaAspLeuValSerLysSerAla 50

||||| ||||||| ||||||| ||||||| ||||||| |||||||

112 ACATGTTGGGAAGGATCTACTTTGGCAGATGGTGTCCAAAAGTGCA 161

||||| ||||||| ||||||| ||||||| ||||||| |||||||

51 GlnTyrCysTyrValAspArgAsnProValGlyLeuMetLeuLeu 67

||||| ||||||| ||||||| ||||||| ||||||| |||||||

162 AACTACTGCCACACATCTCAGGGAGACCCGATTGGCTTAATCTGCTGG 211

67 rGluValAlaLeuGlyAspMetTyrGluLeuLysLysAlaThrSerMetA 84

||||| ||||||| ||||||| ||||||| ||||||| |||||||

212 AGAGGTGGCCCTTGGAAACATGTATGAACATCAAGCATGCTTCACATATCA 261

||||| ||||||| ||||||| ||||||| ||||||| |||||||

84 sPLysProProArgGlyLysHisSerThrLysGlyLeuGlyLysThrVal 100

||||| ||||||| ||||||| ||||||| ||||||| |||||||

262 GCAAGTTACCCAGGGCAAGCACAGTGTCAAAGGTTGGGAAAACCAACC 311

||||| ||||||| ||||||| ||||||| ||||||| |||||||

101 ProLeuSerGluPheValLysTyrArgAspValValValProCy 117

||||| ||||||| ||||||| ||||||| ||||||| |||||||

312 CCTGACCCCTTCGGCCACATCAACC...CTGGAGGGGTAGAGGTTCCACT 358

||||| ||||||| ||||||| ||||||| ||||||| |||||||

117 sGlyLysProValProSerSerIleArgSerSerGluLeuMetTyrAsnG 134

||||| ||||||| ||||||| ||||||| ||||||| |||||||

359 GGAACAGGGATCCCATCTGCTCAACGACACCTGCCTGCTGTATATG 408

||||| ||||||| ||||||| ||||||| ||||||| |||||||

134 LuTyrIleValTyrAsnThrSerGlnValLysMetGlnPheLeuLys 150

||||| ||||||| ||||||| ||||||| ||||||| |||||||

409 AGTACATTGTCTACGACATTGCTCAGGTGAATCTCAAAATACCTGCTGAAA 458

||||| ||||||| ||||||| ||||||| ||||||| |||||||

151 ValArgPheHisLys 156

||||| ||||||| ||||||| ||||||| ||||||| |||||||

459 CTCAAGTTCAAATTTAAG 476

||||| ||||||| ||||||| ||||||| ||||||| |||||||

seq_name: gb_est1:BE304535

seq_documentation_block:

LOCUS BE304535 571 bp mRNA EST 13-JUL-2000
 DEFINITION 601105502F1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:2988216 5',
 mRNA sequence.

ACCESSION BE304535

VERSION BE304535.1 GI:9175756

KEYWORDS EST.

SOURCE human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 571)

AUTHORS NIH-MGC http://mgc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

JOURNAL

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: NIH Intramural Sequencing Center

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov

Plate: LLCM77 row: c column: 01

High quality sequence stop: 567.

FEATURES

source

1..571

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:2988216"

/clone_lib="NIH_MGC_15"

/tissue_type="adenocarcinoma cell line"

/lab_host="DH10B (phage-resistant)"

/note="Organ: colon; Vector: pOTB7; Site_1: XhoI; Site_2:

EcoRI; cDNA made by oligo-dT priming. Directionally

cloned into EcoRI/XhoI sites using the following 5'

adaptor: GGCACG(G). Size selected >500bp for average

insert size 1.8kb. Library constructed by Ling Hong in

the laboratory of Gerald M. Rubin (University of

California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies)"

147 a 140 c 140 g 144 t

BASE COUNT

ORIGIN

alignment_scores:

Quality: 517.50 Length: 156
Ratio: 3.777 Gaps: 1
Percent Similarity: 87.821 Percent Identity: 62.179

alignment_block:

US-09-236-995D-4 x BE304535 ..

Align seg 1/1 to: BE304535 from: 1 to: 571

```

1 AsnLysMetLeuLeuTyrPheHisGlySerArgLeuThrAsnPheValGlyI1 17
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
40 AACCCGAAGATTGCTGTGGCAGCGGTCCAGGACCACCAACTTTGCTGGGAT 89
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
17 eLeuSerGlnGlyLeuArgIleAlaProGluAlaProValThrGlyT 34
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
90 CCTGTCAGGGTCTCGGATAGCCCCCGCTGACATGGCTCCAGAGTGC 139
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
34 yrMetPheGlyLysGlyLeuTyrPheAlaAspLeuValSerLysSerAla 50
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
140 ACATGTTGGTAAAGGATCTATTTCGCTGACATGGCTCCAGAGTGC 189
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
51 GlnTyrCysTyrValAspArgAsnProValGlyLeuMetLeuLeuSe 67
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
190 AACTACTGCCATACGCTCCAGGAGACCCCAATAGGCTTAATCCTGTTGG 239
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
67 rGluValAlaLeuGlyAspMetTyrGluLeuLysAlaThrSerMetA 84
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
240 AGAAGTTGCCCTTGGAAACATGTATGAACCTGAACGACGCTTCACATATCA 289
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
84 sPLysProProArgGlyLysHisSerThrLysGlyLeuGlyLysThrVal 100
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
290 GCAAGTTACCCCAAGGCAAGCACAGTGTCAAAAGTTTGGGCAAACTACC 339
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
101 ProLeuGluSerGluPheValLysTrpArgAspValValValProCy 117
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
340 CTGATCCTTCAGGTACATAGT...CTGGATGGTGTAGACGTTCTCT 386
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
117 sGlyLysProValProSerSerIleArgSerGluLeuMetTyrAsnG 134
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
387 TGGACCGGGATTTCATCTGTGTGTGAATGACACCTCTCTACTATATAAG 436
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
134 LuTyrIleValTyrAsnThrSerGlnValLysMetGlnPheLeuLeuLys 150
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
437 AGTACATTGTCTATGATATTGCTCAGGTAAATCTGAAGTATCTGCTGAAA 486
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
151 ValArgPheHisLysLys 156
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
487 CTGAATTCATATTTTAAG 504
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

```

seq_name: gb_estl:AA397988

seq_documentation_block:

LOCUS AA397988 579 bp mRNA EST 12-AUG-1997
DEFINITION zt86c12.r1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:729238
5' similar to gb:J03473 NAD(+) ADP-RIBOSYLTRANSFERASE (HUMAN);
mRNA sequence.
ACCESSION AA397988
VERSION AA397988.1 GI:2050655
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 579)
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B.,
Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie,
T., Waterston,R. and Wilson,R.
TITLE WashU-Merck EST Project 1997
JOURNAL Unpublished (1997)
CONTACT: Wilson RK
COMMENT Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.

Insert Length: 1320 Std Error: 0.00

Seq primer: -28ml3 rev2 ET from Amersham

High quality sequence stop: 427.

FEATURES

source

1..579 Location/Qualifiers

/organism="Homo sapiens"

/db_xref="GDB:592576"

/db_xref="taxon:9606"

/clone="IMAGE:729238"

/clone.lib="Soares_testis_NHT"

/sex="male"

/lab_host="DH10B"

/note="vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech Laboratories
, Inc. and primed with a Not I - oligo(dT) primer [5',
TGTTACCAATCTGAAGTGGGAGCGCGCCCAATTTTTTTTTTTT 3'].

Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 148 a 138 c 144 g 149 t

ORIGIN

alignment_scores:

Quality: 517.50 Length: 156
Ratio: 3.777 Gaps: 1
Percent Similarity: 87.821 Percent Identity: 62.179

alignment_block:

US-09-236-995D-4 x AA397988 ..

Align seg 1/1 to: AA397988 from: 1 to: 579

```

1 AsnLysMetLeuLeuTyrPheHisGlySerArgLeuThrAsnPheValGlyI1 17
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
43 AACCCGAAGATTGCTGTGGCAGCGGTCCAGGACCACCAACTTTGCTGGGAT 92
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
17 eLeuSerGlnGlyLeuArgIleAlaProGluAlaProValThrGlyT 34
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
93 CCTGTCAGGGTCTTCGGATAGCCCCCGCTGAAGCGCCCGTGACAGGCT 142
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
34 yrMetPheGlyLysGlyLeuTyrPheAlaAspLeuValSerLysSerAla 50
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
143 ACATGTTGGTAAAGGATCTATTTCGCTGACATGGCTCCAGAGTGC 192
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
51 GlnTyrCysTyrValAspArgAsnProValGlyLeuMetLeuLeuSe 67
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
193 AACTACTGCCATACGCTCTCAGGAGACCCCAATAGGCTTAATCCTGTTGG 242
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
67 rGluValAlaLeuGlyAspMetTyrGluLeuLysLysAlaThrSerMetA 84
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
243 AGAAGTTGCCCTTGGAAACATGTATGAACCTGAAGCAGCAGCTTCACATATCA 292
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
84 sPLysProProArgGlyLysHisSerThrLysGlyLeuGlyLysThrVal 100
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
293 GCAAGTTACCCCAAGGCAAGCACAGTGTCAAAAGTTTGGGCAAACTACC 342
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
101 ProLeuGluSerGluPheValLysTrpArgAspValValValProCy 117
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
343 CCTGATCCTTCAGCTAACATAGT...CTGGATGGTGTAGACGTTCTCT 389
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
117 sGlyLysProValProSerSerIleArgSerGluLeuMetTyrAsnG 134
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
390 TGGGACCGGGATTTTCATCTGGTGAATGACACCTCTCTACTATATAAG 439
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

```

134 luTyrileValtyrAsnThrSerGlnVallyMetGlnPheLeuLeuLys 150
 |||||
 440 AGTACATTGCTATGATATTCCTCAGGTAAATCTGAAGTATCTGCTGAAA 489
 151 ValArgPheHisLysLys 156
 |||||
 490 CTGAATTCATTTTAAAG 507

seq_name: gb_est1:AA401836

seq_documentation_block: 669 bp mRNA EST 16-MAY-1997
 LOCUS AA401836
 DEFINITION zv66912.r1 Soares_total_fetus_Nb2HF8_9w Homo sapiens cDNA clone
 IMAGE:758591 5' similar to gb:J03473 NAD(+) ADP-RIBOSYLTRANSFERASE
 (HUMAN); mRNA sequence.

ACCESSION AA401836
 VERSION AA401836.1 GI:2055855
 KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
 Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B.,
 Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie,
 T., Waterston,R. and Wilson,R.

TITLE WashU-Merck EST Project 1997

JOURNAL Unpublished (1997)

COMMENT Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available royalty-free through LNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: -28ml3 rev2 Et from Amersham

High quality sequence stop: 493.

FEATURES

source

1..669
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:758591"
 /clone_lib="Soares_total_fetus_Nb2HF8_9w"
 /dev_stage="8-9 weeks"
 /lab_host="DH10B"
 /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
 polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
 was prepared from mRNA obtained from pooled 8-9 week
 (total) fetus material with a Not I - oligo(dT) primer [5'
 TGTTACCAATCGAATGGAGCGCGCTTAATTTTTTTTTTTTTTTT 3'].

Double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT7T3 vector. Library
 went through one round of normalization, and was
 constructed by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT 191 a 152 c 157 g 169 t
 ORIGIN

alignment_scores:

Quality: 517.50 Length: 156
 Ratio: 3.777 Gaps: 1
 Percent Similarity: 87.821 Percent Identity: 62.179

alignment_block:

US-09-236-995D-4 x AA401836 ..

Align seg 1/1 to: AA401836 from: 1 to: 669

1 AsnLysMetLeuLeuTrpHisGlySerArgLeuThrAsnPheValGlyI1 17

|||||
 200 AACCGAAGATTGCTGGCACGGGTCCAGGACCACCACTTTGCTGGGAT 249
 17 eLeuSerGlnGlyLeuArgIleAlaProGluAlaProGluAlaProValThrGlyT 34
 |||||
 250 CCTGTCCACAGGTCCTCGGATAGCCCGCCGCTGAAGCCCGCTGACAGGCT 299
 34 yrMetPheGlyLysGlyLeuTyrPheAlaAspLeuValSerLysSerAla 50
 |||||
 300 ACATGTTTGGTAAAGGATCTATTTTCGCTGACATGGTCTCCAAGAGTGCC 349
 51 GlnTyrCysTyrValAspArgAsnProValGlyLeuMetLeuLeu 67
 |||||
 350 AACTACTGCATACGTCCTCAGGAGACCAATAGGCTTAATCCTGTTGGG 399
 67 rGluValAlaLeuGlyAspMetTyrGluLeuLysLysAlaThrSerMetA 84
 |||||
 400 AGAAGTTGCCCTTGGAAACATGATGAACAGACAGCGCTTCACATATCA 449
 84 sPlysProProArgGlyLysHisSerThrLysGlyLeuGlyLysThrVal 100
 |||||
 450 GCAAGTTACCCCAAGGCGCAAGCAGGTGTCAAAGTTTGGGCAAACTACC 499
 101 ProLeuGluSerGluPheVallyLysTrpArgAspValValProCy 117
 |||||
 500 CCGTATCCTCTCAGCTAACATTAGT...CTGGATGGTGTAGACGTTCCCTCT 546
 117 sGlyLysProValProSerSerIleArgSerSerGluLeuMetTyrAsnG 134
 |||||
 547 TGGGACCGGGATTTCATCTGGTGTGAATGACACCTCTCTACTATATAACG 596
 134 luTyrileValtyrAsnThrSerGlnVallyMetGlnPheLeuLeuLys 150
 |||||
 597 AGTACATTGCTATGATATTGCTCAGGTAATCTGAAGTATCTGCTGAAA 646
 151 ValArgPheHisLysLys 156
 |||||
 647 CTGAAATTCATTTTAAAG 664

seq_name: gb_est2:BI093436

seq_documentation_block:

LOCUS BI093436 916 bp mRNA EST 20-JUN-2001
 DEFINITION 602859629F1 NIH_MGC_10 Homo sapiens cDNA clone IMAGE:5000911 5',
 mRNA sequence.

ACCESSION BI093436

VERSION BI093436.1 GI:14511766

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS NIH-MGC http://mgc.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: Incyte Genomics, Inc.

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LAM11033 row: i column: 08

High quality sequence stop: 837.

FEATURES

source

1..916
 /organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5000911"

/clone_lib="NIH_MGC_10"

BASE COUNT 248 a 209 c 230 g 229 t
ORIGIN

/cell_line="MGC36"
/lab_host="DH10B"
/note="Organ: cervix; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.5 kb. Library prepared by Life
Technologies."

alignment_scores:
Quality: 517.50 Length: 156
Ratio: 3.777 Gaps: 1
Percent Similarity: 87.821 Percent Identity: 62.179

alignment_block:
US-09-236-995D-4 x BI093436 ..
Align seg 1/1 to: BI093436 from: 1 to: 916

1 AsnLysMetLeuThrHisGlySerArgLeuThrAsnPheValGlyI 17
|||||
185 AACCGAAGATTGCTGGCAGCGGTCCAGGACCACCACTTGTGGGAT 234
17 eLeuSerGlnGlyLeuArgIleAlaProGluAlaProValThrGlyT 34
|||||
235 CCTGTCCAGGGTCTTCGGATAGCCGCGCTGAAGCGCGGTGACAGGCT 284
34 yMetPheGlyLysGlyLeuThrPheAlaAspLeuValSerLysSerAla 50
|||||
285 ACATGTTTGTAAAGGATCTATTTCCGTCACATGGTCCAGAGTGCC 334
51 GlnTyrCystyrValAspArgAsnProValGlyLeuMetLeuLeuSe 67
:|||||
335 AACTACTGCCATACGCTCAGGAGACCAATAGGCTTAATCTGTGGG 384
67 xGluValAlaLeuGlyAspMetTyrGluLeuLysLysAlaThrSerMetA 84
|||||
385 AGAAGTTGCCCTTGGAACATGTATGAACCTGAAGCAGCTTCACATCA 434
84 sPLysProArgGlyLysHisSerThrLysGlyLeuGlyLysThrVal 100
:|||||
435 GCAAGTTACCAAGGCGAACACAGTGTCAAGGTTTGGGCAAACTACC 484
101 ProLeuGluSerGluPheValLysTrpArgAspValValValProCy 117
|||
485 CCTGATCCTTCAGCTAACATTAGT...CTGGATGGGTAGAGCTTCTCT 531
117 sGlyLysProValProSerSerIleArgSerSerGluLeuMetTyrAsnG 134
|||
532 TGGGACCGGGATTCAFTCTGGTGTGAATGACACCTCTCTACTATAAACG 581
134 luTyrIleValTyrAsnThrSerGlnValLysMetGlnPheLeuLys 150
|||||
582 AGTACATTGTATGATATGCTCAGGTAAATCTGAAGTATCTGCTGAAA 631
151 ValArgPheHisLys 156
:|||||
632 CTGAATTCATTTAAAG 649

seq_name: gb_est2:BF793705

seq_documentation_block:
LOCUS BF793705 952 bp mRNA 12-JAN-2001
DEFINITION 60225121F1 NIH_MGC_84 Homo sapiens cDNA clone IMAGE.4347365 5',
mRNA sequence.
ACCESSION BF793705
VERSION BF793705.1 GI:12098759
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 952)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9970 row: j column: 06
High quality sequence stop: 672.

FEATURES

source
1..952
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4347365"
/clone_lib="NIH_MGC_84"
/tissue_type="adrenal cortex carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: adrenal gland; Vector: pCMV-SPORT6; Site 1:
NotI; Site 2: SalI; Cloned unidirectionally; Oligo-dT
primed. Average insert size 1.229 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT 231 a 215 c 244 g 262 t
ORIGIN

alignment_scores:

Quality: 517.50 Length: 156
Ratio: 3.777 Gaps: 1
Percent Similarity: 87.821 Percent Identity: 62.179

alignment_block:

US-09-236-995D-4 x BF793705 ..
Align seg 1/1 to: BF793705 from: 1 to: 952
1 AsnLysMetLeuThrHisGlySerArgLeuThrAsnPheValGlyI 17
|||||
52 AACCGAAGATTGCTGGCAGCGGTCCAGGACCACCACTTGTGGGAT 101
17 eLeuSerGlnGlyLeuArgIleAlaProGluAlaProValThrGlyT 34
|||||
102 CCTGTCCAGGGTCTTCGGATAGCCGCGCTGAAGCGCGGTGACAGGCT 151
34 yMetPheGlyLysGlyLeuThrPheAlaAspLeuValSerLysSerAla 50
|||||
152 ACATGTTTGTAAAGGATCTATTTCCGTCACATGGTCTCCAAGAGTGCC 201
51 GlnTyrCystyrValAspArgAsnProValGlyLeuMetLeuLeuSe 67
:|||||
202 AACTACTGCCATACGCTCAGGAGACCAATAGGCTTAATCTGTGGG 251
67 rGluValAlaLeuGlyAspMetTyrGluLeuLysLysAlaThrSerMetA 84
|||||
252 AGAAGTTGCCCTTGGAACATGTATGAACCTGAAGCAGCTTCACATCA 301
84 sPLysProArgGlyLysHisSerThrLysGlyLeuGlyLysThrVal 100
:|||||
302 GCAAGTTACCAAGGCGAACACAGTGTCAAGGTTTGGGCAAACTACC 351
101 ProLeuGluSerGluPheValLysTrpArgAspValValValProCy 117
|||
352 CCTGATCCTTCAGCTAACATTAGT...CTGGATGGGTAGAGCTTCTCT 398
117 sGlyLysProValProSerSerIleArgSerSerGluLeuMetTyrAsnG 134
|||
399 TGGGACCGGGATTTTCATCTGGTGTGAATGACACCTCTCTACTATAAACG 448

```
134 lutyriLeValTyrAnThrSerGlnValLysMetGlnPheLeuLys 150
|||||
449 AGTACATTGCTATGATATGCTCAGGTAATCTGAAGTATCTGCTGAAA 498
151 ValArgPheHisLys 156
499 CTGAATTCATTTAAG 516
seq_name: gb_est2:BE911127
seq_documentation_block:
LOCUS BE911127 587 bp mRNA EST 29-SEP-2000
DEFINITION 601662539F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:3962696 5',
mRNA sequence.
ACCESSION BE911127
VERSION BE911127.1 GI:10408111
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 587)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM9129 row: n column: 09
High quality sequence stop: 559.
FEATURES
Location/Qualifiers
1..587
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:3962696"
/clone_lib="NCI_CGAP_Mam1"
/tissue_type="tumor, biopsy sample"
/dev_host="3 months, virgin"
/lab_host="DH10B"
/Note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"
BASE COUNT 149 a 163 c 150 g 125 t
ORIGIN
alignment_scores:
Quality: 513.50 Length: 156
Ratio: 3.776 Gaps: 1
Percent Similarity: 87.179 Percent Identity: 61.538
alignment_block:
US-09-236-995D-4 x BE911127 ..
Align seg 1/1 to: BE911127 from: 1 to: 587
1 AsnLysMetLeuLeuTrpHisGlySerArgLeuThrAsnPheValGlyTl 17
|||||
25 AACCGGAGGCTGTGTGGCAGCGCTCCAGGACCACTTTGCTGGCAT 74
17 eLeuSerGlnGlyLeuArgTleAlaProGluAlaProValThrGlyT 34
|||||
75 CCTGTCCAGGGTCTGGGGATAGCCCCCACTGAGCGCTGTGACAGCT 124
```

```
34 YrMetPheGlyLysGlyLeuTyrPheAlaAspLeuValSerLysSerAla 50
|||||
125 ACATGTTGGAAAGGATCTACTTTGCCGACATGCTGTCACAAAGTCCA 174
51 GlnTyrCysTyrValAspArgAsnProValGlyLeuMetLeuLeuSe 67
|||||
175 AACTACTGCCACACATCTCAGGAGACCCGATTGGCTTAATATCTGCTGG 224
67 rGluValAlaLeuGlyAspMetTyrGluLeuLysLysAlaThrSerMetA 84
|||||
225 AGAGGTGCCCTTGGAAACATGATGAACCTCAAGCATGCTTCACATATCA 274
84 sPlyProProArgGlyLysHisSerThrLysGlyLeuGlyLysThrVal 100
|||||
275 GCAAGTTTACCACCAAGGGCAAGCAGTGTCAAAGGTTTGGGAAAACACC 324
101 ProLeuGluSerGluPheValLysTrpArgAspValValProCy 117
|||||
325 CCTGACCTTTCGGCCACATCACC...CTGGAGGGGTGTAGAGGTTCCACT 371
117 sGlyLysProValProSerSerIleArgSerSerGluLeuMetTyrAsnG 134
|||||
372 GGAACAGGATCCCATCTCGTGTCAACGACACCTGCCTCTGTATCATG 421
134 lutyriLeValTyrAnThrSerGlnValLysMetGlnPheLeuLys 150
|||||
422 AGTACATTGCTACGACATGCTCAGTGAATCTCAAAATACCTGCTGAAA 471
151 ValArgPheHisLys 156
472 CTCAAGTTCAATTTAAG 489
seq_name: gb_est1:BE382739
seq_documentation_block:
LOCUS BE382739 709 bp mRNA EST 21-JUL-2000
DEFINITION 601297572F1 NIH_MGC_19 Homo sapiens cDNA clone IMAGE:3627881 5',
mRNA sequence.
ACCESSION BE382739
VERSION BE382739.1 GI:9328104
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 709)
NIH-MGC http://mgi.nci.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: L1CM311 row: o column: 18
High quality sequence stop: 646.
FEATURES
Location/Qualifiers
1..709
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3627881"
/clone_lib="NIH_MGC_19"
/tissue_type="neuroblastoma"
/lab_host="DH10B (phage-resistant)"
/Note="Organ: brain; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(C). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
```

(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

BASE COUNT 178 a 167 c 181 g 183 t
ORIGIN

alignment_scores:
Quality: 510.50 Length: 156
Ratio: 3.754 Gaps: 1
Percent Similarity: 87.179 Percent Identity: 61.538

alignment_block:
US-09-236-995D-4 x BE382739 ..

Align seg 1/1 to: BE382739 from: 1 to: 709

1 AsnLysMetLeuLeuTyrPheHisGlySerArgLeuThrAsnPheValGlyI 17
12 AACCGAAGATTGCTGGCAGCGGTCCAGGACCAACCTTGGTGGGAT 61
17 eLeuSerGlnGlyLeuArgIleAlaProProGluAlaProValThrGlyT 34
62 CCTGTCAGGATCTCGGATAGCCCGCTGAAGCGCCGTCGACAGGCT 111
34 yrMetPheGlyGlyLeuTyrPheAlaAspLeuValSerLysSerAla 50
112 ACATGTTGGTAAAGGATCTATTTCGCTGACATGCTCCAAAGAGTGCC 161
51 GlnTyrCysTyrValAspArgAsnAspProValGlyLeuMetLeuLeuSe 67
162 AACTACTGCGCATAGCTCTAGGAGAGCCCAATAGCTTAATCTGTGGG 211
67 rGluValAlaLeuGlyAspMetTyrGluLeuLysLysAlaThrSerMetA 84
212 AGAAGTTCCTTGGAAACATGATGAACCTGAAGCAGCTTCACATATCA 261
84 spLysProProArgGlyLysHisSerThrLysGlyLeuGlyLysThrVal 100
262 GCAAGTTACCAACCAAGGCAAGCAGGTCAAGGTTTGGGCAAACTACC 311
101 ProLeuGluSerGluPheValLysTyrArgAspAspValValProCy 117
312 CTGATCTTCACTAACATTAGT...CTGGATGGGTAGACCTTCTCT 358
117 sGlyLysProValProSerSerLysArgSerGluLeuMetTyrAsnG 134
359 TGGACCGGATTTCTATCTGCTGCTGAATGACACCTCTCTACTATATA 408
134 lutyrlleValTyrAsnThrSerGlnValLysMetGlnPheLeuLys 150
409 AGTACATTGCTATGATATTGCTCAGGTAATCTGAAGTATCTGCTGAAA 458
151 ValArgPheHisLys 156
459 CTGAATTCATTTTAAAG 476

seq_name: gb_est2:BI079902

seq_documentation_block:
LOCUS BI079902 725 bp mRNA EST 20-JUN-2001
DEFINITION 602876096F1 NCI_CGAP_Mam2 Mus musculus cDNA clone IMAGE:5007983 5',
mRNA sequence.
ACCESSION BI079902
VERSION BI079902.1 GI:14498232
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 725)
NH-MGC <http://mgc.nci.nih.gov/>
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM11051 row: o column: 24
High quality sequence stop: 664.

FEATURES

source
1..725
/organism="Mus musculus"
/strain="FVB/N-3"
/db_xref="taxon:10090"
/clone_image="5007983"
/clone_lib="NCI_CGAP_Mam2"
/tissue_type="tumor, biopsy sample"
/dev_stage="5 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site: 1: Sali;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"
BASE COUNT 177 a 204 c 188 g 156 t
ORIGIN

alignment_scores:

Quality: 509.50 Length: 153
Ratio: 3.802 Gaps: 1
Percent Similarity: 87.582 Percent Identity: 62.745

alignment_block:

US-09-236-995D-4 x BI079902 ..
Align seg 1/1 to: BI079902 from: 1 to: 725

4 LeuLeuTyrPheHisGlySerArgLeuThrAsnPheValGlyIleLeuSerG 20
3 CTGCTGTGGCAGCGGTCCAGGACCACTTGTGTCGATCCTGTGCA 52
20 nGlyLeuArgIleAlaProGluAlaProValThrGlyTyrMetPheG 37
53 GGCTCTGGGATAGCCCACTGAAGCCCTGTGACAGCTACATGTTG 102
37 lYlYSGLyLeuTyrPheAlaAspLeuValSerLysSerAlaGlnTyrCys 53
103 GGAAGGGATCTACTTCCGACATGGTGTCCAAAGTGCAAACTACTGC 152
54 TyrValAspArgAsnAspProValGlyLeuMetLeuLeuSerGluValAl 70
153 CACACATCTCAGGAGACCGGATTTGGCTTAATCTGCTGGAGAGGTGC 202
70 aLeuGlyAspMetTyrGluLeuLysLysAlaThrSerMetAspLysProp 87
203 CCTTGGAAACATGATGAACTCAAGCATGCTTCACATATCAGCAAGTAC 252
87 roArgGlyLysHisSerThrLysGlyLeuGlyLysThrValProLeuGlu 103
253 CCAAGGGCAAGCACAGTGTCAAAGGTTTGGGAAAACCAACCCCTGACCT 302
104 SerGluPheValLysTyrArgAspValValValProCysGlyLysPr 120
303 TCGGCCAGCATCACC...CTGGAGGGTGTAGAGGTTCCTCTGGGAACAGG 349
120 oValProSerSerIleArgSerGluLeuMetTyrAsnGluTyrIleV 137
350 GATCCCATCTGCTGACAGCACCTGCTGCTGTATATGAGTACATG 399
137 alTyrAsnThrSerGlnValLysMetGlnPheLeuLeuLysValArgPhe 153

400 TCTACGACATTGCTCAGTGGAATCTCAAAATCCTGCTGCTGAAGTTC 449

154 HisHisLys 156

450 AATCTTAAG 458

seq_name: gb_est2:BG675913

seq_documentation_block: 928 bp mRNA EST 01-MAY-2001
LOCUS BG675913 602622263f1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4747574 5',
DEFINITION mRNA sequence.

ACCESSION BG675913

VERSION BG675913.1 GI:13907309

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 928)

NTI-MGC http://mgc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: Cgapbs-r@mail.nih.gov

Tissue Procurement: James Cleaver, M.D.

cDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone Distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LIA10597 row: m column: 15

High quality sequence stop: 816.

Location/Qualifiers

1..928

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:4747574"

/clone_lib="NCI_CGAP_Skn4"

/tissue_type="squamous cell carcinoma"

/lab_host="DH10B (TI phage-resistant)"

/note="Organ: skin; Vector: pCMV-SPORT6; Site: 1: NotI;

Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.

Average insert size 1.5kb. Library constructed by Life

Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT 247 a 209 c 225 g 247 t

ORIGIN

alignment_scores: Quality: 503.50 Length: 153

Ratio: 3.757 Gaps: 1

Percent Similarity: 87.582 Percent Identity: 62.092

alignment_block:

US-09-236-995D-4 x BG675913

Align seg 1/1 to: BG675913 from: 1 to: 928

4 LeuLeuTrpHisGlySerArgLeuThrAsnPhValGlyLeuLeuSergl 20

|||||

6 TTGCTGTGGCAGCGGTCCAGGACCACTTTGCTGGGATCTCTGTCCTCA 55

|||||

20 nGlyLeuArgIleAlaProGluAlaProValThrGlyTyrMetPheG 37

|||||

56 GGATCTTGGATAGCCCGCTGAAGCCCGGTGACAGGCTACATGTTTG 105

|||||

37 lYlYsGlyLeuTyrPheAlaAspLeuValSerLysSerAlaGlnTyrCys 53

|||||

106 GTRAAAGGATCTATTTCGCTGACATGCTCCCAAGAGTGCCAACTACTGC 155

|||||

54 TyrValAspArgAsnAnProValGlyLeuMetLeuLeuSerGluValAl 70

156 CATACGCTCTCAGGAGACCAATAGCTTAATCCTGTTGGGAGAAGTGC 205

70 aLeuGlyAspMetTyrGluLeuLysLysAlaThrSerMetAspLysProp 87

|||||

206 CTTTGAACATCATGAACTGAAGCAGCTTCACATATCAGCAAGTTAC 255

|||||

87 roArgGlyLysHisSerThrLysGlyLeuGlyLysThrValProLeuGlu 103

|||||

256 CCAAGGCAAGCAGACAGTGTCAAGAGTTTGGCAAAACTACCCCTGATCCT 305

|||||

104 SerGluPheValLysTrpArgAspValValValProCysGlyLysPr 120

|||||

306 TCAGTAAACATTAGT...CTGGATGGTGTAGACGTTCTCTTGGGACCGG 352

|||||

120 ovalProSerSerIleArgSerSerGluLeuMetTyrAsnGluTyrIleV 137

|||||

353 GATTTCATCTGGTGTGAATGACACCTCTCTACTATATACGAGTATATG 402

|||||

137 alTyrAsnThrSerGlnValLysMetGlnPheLeuLeuLysValArgPhe 153

|||||

403 TCTATGATATTGCTCAGGTAATCTGAAGTATCTGCTGAAACTGAAATTC 452

|||||

154 HisHisLys 156

453 AATTTTAAG 461

seq_name: gb_est2:BG042229

seq_documentation_block: 502 bp mRNA EST 31-JUL-2001

LOCUS BG042229 su93c12.y1 Gm-cl055 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:

DEFINITION Gm-cl055-2064 5' similar to TR:Q9ZP54 Q9ZP54 POLY (ADP-RIBOSE)

POLYMERASE ;, mRNA sequence.

ACCESSION BG042229

VERSION BG042229.1 GI:12488706

KEYWORDS EST.

SOURCE soybean.

ORGANISM Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;

Glycine.

1 (bases 1 to 502)

Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Corveill, V., Khanna

, A., Boll, B., Marra, M., Hillier, L., Kucaba, F., Martin, J., Beck, C.,

Wyllie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers

, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk

, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann

, R., Waterston, R. and Wilson, R.

Public Soybean EST Project

Unpublished (1999)

Contact: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available through: Genome Systems, Inc. 4633 World

Parkway Circle St. Louis, Missouri 63134 For further information

call: (800) 430-0030 or (314) 427-3222 FAX:(888) 919-3324 or (314)

427-3324 or contact: clones@genomesystems.com or

info@genomesystems.com web site: www.genomesystems.com

Insert Length: 429 Std Error: 0.00

High quality sequence stop: 419.

Location/Qualifiers

1..502

/organism="Glycine max"

/db_xref="taxon:3847"

/clone="GENOME SYSTEMS CLONE ID: Gm-cl055-2064"

/clone_lib="Gm-cl055"

/tissue_type="Mature seed pods, greenhouse grown"

FEATURES

SOURCE

```

/lab_host="DH10B"
/notes="Vector: pBluescript II SK+; Site_1: EcoRI; Site_2:
XhoI; The cDNA library was constructed from mRNA isolated
from mature seed pods of greenhouse grown plants prior to
senescence for the cultivar KPL. Complementary DNA was
synthesized from mRNA using a primer consisting of a
poly(dT) sequence with a XhoI restriction site. EcoRI
adapters were ligated to the blunt-ended cDNA fragments
followed by XhoI digestion. The cDNA fragments were
directionally cloned into the EcoRI-XhoI restriction site
of the pBluescript vector. The ligated cDNA fragments
were transformed into DH10B host cells (GibcoBRL). This
library was constructed in the laboratory of Dr. Randy
Shoemaker."

```

```

BASE COUNT      150 a      80 c      119 g      153 t
ORIGIN

```

alignment_scores:

```

Quality: 493.00      Length: 110
Ratio: 4.695          Gaps: 0
Percent similarity: 95.455      Percent Identity: 81.818

```

alignment_block:

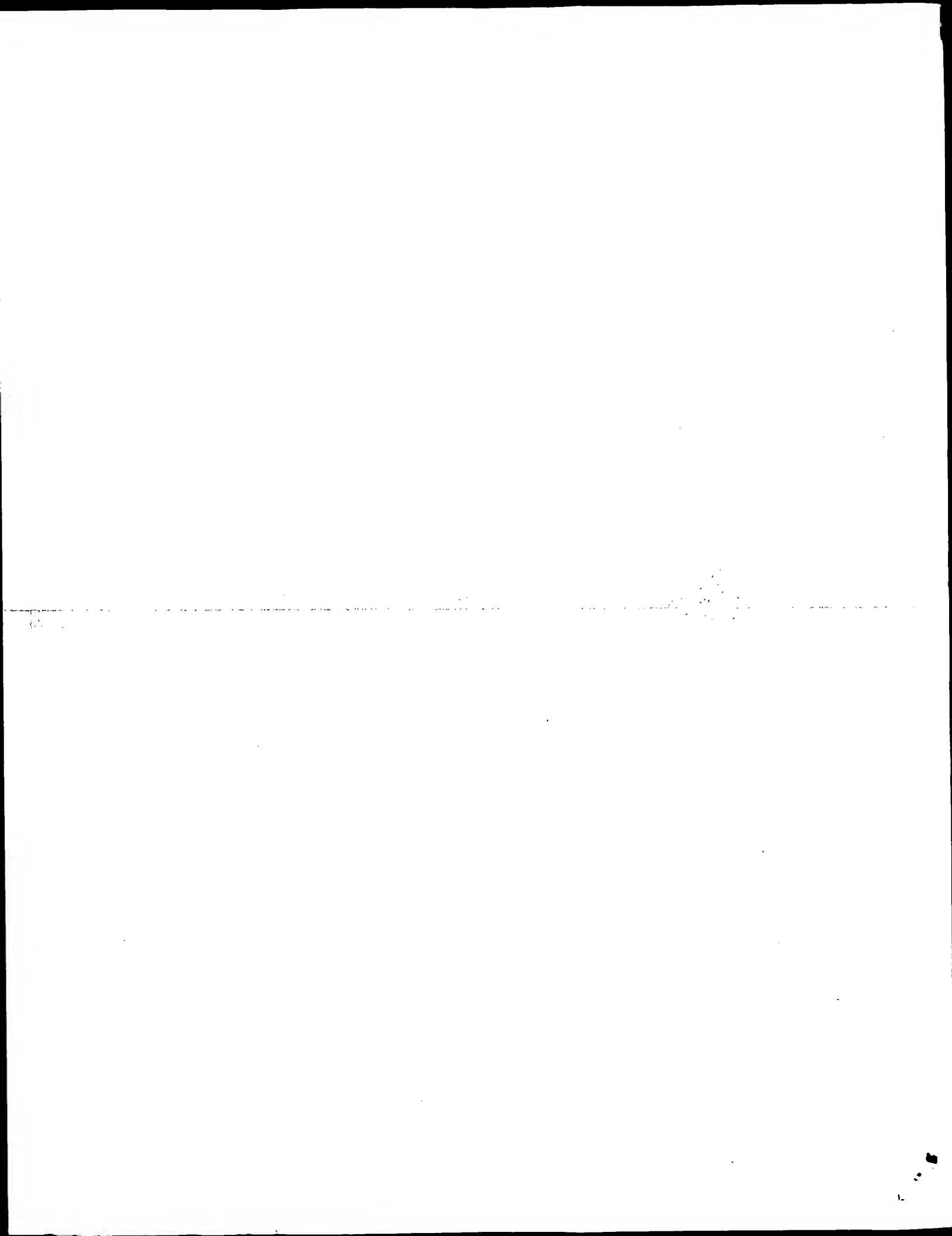
```
US-09-236-995D-4 x BG042229 ..
```

```
Align seg 1/1 to: BG042229 from: 1 to: 502
```

```

48  LysSerAlaGlnTyrCysTyrValAspArgAsnAsnProValGlyLeuMe 64
|||||
4  AAAAGTCTCAGTATTGCITCACTGATAAGAAAAATCCTGTGGTCTAAT 53
|||||
64  tLeuLeuSerGluValAlaLeuGlyAspMetTyrGluLeuLysLysAla 81
|||||
54  GCCTTTGAGTGAAGTTGCGCTTGGAATGCTATGAGCTCAAGAAAGCTA 103
|||||
81  hrSerMetAspLysProProArgGlyLysHisSerThrLysGlyLeuGly 97
|||||
104 AGTATATGGATAAACCTCCGGAAGAAAGCACTCTACTAAAGGACTGGGC 153
|||||
98  LysThrValProLeuGluSerGluPheValLysTrpArgAspValVa 114
|||  ::|||  |||||
154 AAGAAATGCCACAGGATCGGATATGTAAGTGGGGGCAATGTCAC 203
|||||
114 lValProCysGlyLysProValProSerSerIleArgSerSerGluLeuM 131
:|||||
204 TGTTCTCTTGCGCAACCAAGTGCATCAAAATGTCAGAGTTCTTGAGCTCA 253
|||||
131 etTyrAsnGluTyrIleValTyrAsnThrSerGlnValLysMetGlnPhe 147
|||||
254 TGTAATAGATATATTGTTTATTAATGCTCAAGTTAAGATGCAATTC 303
|||||
148 LeuLeuLysValArgPheHisHisLysArg 157
|||||
304 TTATTGAAGGTGAGGTTTCATCACAGAGA 333

```



OM of: US-09-236-995D-2 to: N_Geneseq_1101.* out_format : pfs
 Date: Mar 9, 2002 8:26 PM
 About: Results were produced by the GenCore software, version 4.5,
 Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODE=frame+p2n.model -DEV=xlp
 -O=/cgn2.1/USPTO.spool/US09236995/runat_08032002_113149_25135/app_query.fasta_1.1273
 -DB=N_Geneseq_1101 -OFMT=fastaf -SUFFIX=rng -GAPOPT=12.000
 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPEXT=0.000 -LOOPEXT=0.000
 -GAPOPT=4.500 -GAPEXT=0.050 -XGAPOPT=10.000 -XGAPEXT=0.500
 -GAPOPT=6.000 -FGAPEXT=7.000 -START=1 -MATRIX=bysum62
 -DELOP=6.000 -DELEXT=7.000 -GAPOPT=10.000 -YGAPEXT=0.500
 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
 -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
 -NORM=ext -MINLEN=0 -MAXLEN=200000000
 -USER=US09236995 -CGNL_1_309 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPXY
 -WAIT -THREADS=1

Search information block:

Query: US-09-236-995D-2
 Query length: 982
 Database: N_Geneseq_1101.*
 Database sequences: 930621
 Database length: 428662619
 Search time (sec): 250.160000

score_list:

Sequence	Strd Orig	ZScore	EScore	Len	Documentation
/SIDS2/gcgdata/geneseq/NA1999.DAT:AAH89542		5036.00	8378.50	2949	
/SIDS2/gcgdata/geneseq/NA2000.DAT:AAZ60618		4740.00	7882.92	3212	
/SIDS2/gcgdata/geneseq/NA2000.DAT:AAZ60615		4666.50	7760.11	3211	
/SIDS2/gcgdata/geneseq/NA1996.DAT:AAZ13732		1594.50	2625.10	3793	3.8e-138
/SIDS2/gcgdata/geneseq/NA2001.DAT:AAZ63954		1584.00	2610.08	3045	2.6e-137
/SIDS2/gcgdata/geneseq/NA1999.DAT:AAZ23799		1583.00	2606.55	3580	4.1e-137
/SIDS2/gcgdata/geneseq/NA2001.DAT:AAZ63341		1518.50	2500.06	3200	3.5e-131
/SIDS2/gcgdata/geneseq/NA1995.DAT:AAZ04221		1237.00	2023.80	5345	1.2e-104
/SIDS2/gcgdata/geneseq/NA2001.DAT:AAZ04219		948.50	1549.66	3308	3.1e-78
/SIDS2/gcgdata/geneseq/NA2001.DAT:AAZ63990		898.00	1462.87	3066	2.6e-73
/SIDS2/gcgdata/geneseq/NA2001.DAT:AAZ63996		851.50	1393.77	1592	1.5e-69
/SIDS2/gcgdata/geneseq/NA2000.DAT:AAZ44287		851.50	1392.08	1814	1.8e-69
/SIDS2/gcgdata/geneseq/NA2001.DAT:AAZ44280		851.50	1391.90	1843	1.9e-69
/SIDS2/gcgdata/geneseq/NA2001.DAT:AAZ62090		851.50	1391.90	1843	1.9e-69
/SIDS2/gcgdata/geneseq/NA2000.DAT:AAZ71118		851.50	1391.47	1912	2.0e-69
/SIDS2/gcgdata/geneseq/NA2000.DAT:AAZ60617		851.50	1390.14	2147	2.4e-68
/SIDS2/gcgdata/geneseq/NA2000.DAT:AAZ60616		838.50	1367.65	2295	4.2e-68
/SIDS2/gcgdata/geneseq/NA2001.DAT:AAH14941		838.00	1368.52	1980	3.8e-68
/SIDS2/gcgdata/geneseq/NA2001.DAT:AAZ85320		656.50	1074.78	864	8.7e-52
/SIDS2/gcgdata/geneseq/NA1995.DAT:AAZ04220		650.00	1056.89	1592	8.6e-51
/SIDS2/gcgdata/geneseq/NA2000.DAT:AAZ44290		491.00	790.19	1740	6.1e-36
/SIDS2/gcgdata/geneseq/NA2000.DAT:AAZ44291		490.50	790.41	1587	6.0e-36
/SIDS2/gcgdata/geneseq/NA2000.DAT:AAZ44288		475.00	760.42	2265	2.8e-34
/SIDS2/gcgdata/geneseq/NA2000.DAT:AAZ44289		475.00	760.42	2265	2.8e-34
/SIDS2/gcgdata/geneseq/NA2000.DAT:AAZ80132		398.00	647.13	595	5.7e-28
/SIDS2/gcgdata/geneseq/NA2001.DAT:AAZ85314		253.00	397.34	1143	4.7e-14
/SIDS2/gcgdata/geneseq/NA2000.DAT:AAZ45690		246.50	368.44	5490	1.9e-12
/SIDS2/gcgdata/geneseq/NA2001.DAT:AAZ92076		246.50	368.44	5490	1.9e-12
/SIDS2/gcgdata/geneseq/NA2001.DAT:AAZ85318		225.00	358.96	550	6.4e-12
/SIDS2/gcgdata/geneseq/NA2001.DAT:AAZ85319		225.00	358.96	550	6.4e-12
/SIDS2/gcgdata/geneseq/NA2001.DAT:AAZ60686		211.50	340.21	395	7.1e-11
/SIDS2/gcgdata/geneseq/NA2001.DAT:AAH42823		203.50	326.82	110	4.0e-10
/SIDS2/gcgdata/geneseq/NA1999.DAT:AAZ2491		177.50	258.59	3417	5.9e-06
/SIDS2/gcgdata/geneseq/NA1996.DAT:AAZ39622		173.50	251.91	3417	5.9e-06
/SIDS2/gcgdata/geneseq/NA2001.DAT:AAZ85306		161.50	256.07	416	3.5e-06
/SIDS2/gcgdata/geneseq/NA2001.DAT:AAZ63838		157.50	223.41	3984	0.0002
/SIDS2/gcgdata/geneseq/NA2000.DAT:AAZ66826		157.50	222.98	4134	0.0002
/SIDS2/gcgdata/geneseq/NA2000.DAT:AAZ29627		157.50	222.98	4134	0.0002
/SIDS2/gcgdata/geneseq/NA2000.DAT:AAZ29632		157.50	222.03	4491	0.0003
/SIDS2/gcgdata/geneseq/NA2000.DAT:AAZ29633		157.50	221.61	4657	0.0003

/SIDS2/gcgdata/geneseq/NA2001.DAT:AAH42831 + 151.00 238.55 3.3e-05 415
 /SIDS2/gcgdata/geneseq/NA2000.DAT:AAZ66824 + 135.00 184.95 0.0317 4296
 /SIDS2/gcgdata/geneseq/NA2001.DAT:AAZ85304 - 133.50 207.83 0.0017 472
 /SIDS2/gcgdata/geneseq/NA2000.DAT:AAZ77578 + 133.50 195.80 0.0079 1344
 /SIDS2/gcgdata/geneseq/NA2001.DAT:AAZ63919 + 133.50 186.68 0.0254 2971

seq_name: /SIDS2/gcgdata/geneseq/NA1999.DAT:AAH89542

seq_documentation_block:

ID: AAX89542 standard; cDNA; 2949 bp.

AC AAX89542;

DT 06-OCT-1999 (first entry)

DE Maize poly ADP-ribose polymerase gene.

XX PARP; poly ADP-ribose polymerase; metabolism; maize; transform; ss.

OS Zea mays.

PH Key Location/Qualifiers
 FT CDS 1..2949

FT /*tag= a

FT /product= "PARP"

FT /note= "Poly ADP-ribose polymerase"

FT /transl_except= (pos:1453..1455, aa:Xaa)

FT /transl_except= (pos:1495..1497, aa:Xaa)

FT /transl_except= (pos:1504..1506, aa:Xaa)

FT /transl_except= (pos:1534..1536, aa:Xaa)

FT /transl_except= (pos:1537..1539, aa:Xaa)

FT /transl_except= (pos:1561..1563, aa:Xaa)

FT /transl_except= (pos:1582..1584, aa:Xaa)

FT /transl_except= (pos:1588..1590, aa:Xaa)

FT /transl_except= (pos:1594..1596, aa:Xaa)

FT /transl_except= (pos:1975..1977, aa:Xaa)

FT /transl_except= (pos:2011..2013, aa:Xaa)

FT /transl_except= (pos:2014..2016, aa:Xaa)

FT /transl_except= (pos:2026..2028, aa:Xaa)

FT /transl_except= (pos:2047..2049, aa:Xaa)

FT /transl_except= (pos:2050..2052, aa:Xaa)

FT /transl_except= (pos:2053..2055, aa:Xaa)

FT /transl_except= (pos:2056..2058, aa:Xaa)

FT /transl_except= (pos:2077..2079, aa:Xaa)

FT /transl_except= (pos:2095..2097, aa:Xaa)

FT /transl_except= (pos:2107..2109, aa:Xaa)

FT /transl_except= (pos:2116..2118, aa:Xaa)

FT /transl_except= (pos:2119..2121, aa:Xaa)

FT /transl_except= (pos:2122..2124, aa:Xaa)

FT /transl_except= (pos:2125..2127, aa:Xaa)

FT /transl_except= (pos:2128..2130, aa:Xaa)

FT /note= "Xaa = unknown"

WO9937789-A1.

29-JUL-1999.

26-JAN-1999; 99WO-US01591.

27-JAN-1998; 98US-0072785.

(PION-) PIONEER HI-BRED INT INC.

Mahajan P, Zuo Z;

WPI; 1999-444613/37.

P-PSDB; AAY28464.

New maize poly ADP-ribose polymerase gene useful in transforming plants to alter their metabolic state

Claim 1; Page 30-35; 48pp; English.

CC The sequence is the maize poly ADP-ribose polymerase (PARP) gene. The
CC sequence codes a PARP with 982 amino acids (AA28464). PARP is required
CC in the cell in most cases of DNA repair, recombination, rearrangement
CC and transposition. PARP gene and antisense gene can be used to transform
CC plant cells and alter the metabolic state of the transformed cell. This
CC is useful in enhancing disease resistance in plants and methods of
CC genetic transformation of plants. Plants transformed with either a sense
CC or antisense PARP nucleotide sequence may be utilized to increase
CC transformation frequency in plant cells. The enzyme also plays a role in
CC cellular stress, so may be beneficial for prevention of plant disease or
CC pathogen attack.
XX
SQ Sequence 2949 BP; 895 A; 539 C; 732 G; 756 T; 27 other;

alignment_scores:
Quality: 5036.00 Length: 982
Ratio: 5.251 Gaps: 0
Percent Similarity: 97.658 Percent Identity: 97.658

alignment_block:
US-09-236-995d-2 x AAX89542 ..

Align seg 1/1 to: AAX89542 from: 1 to: 2949

1 MetAlaAlaProProLysAlaLirPlysAlaGluTyrAlaLysSerClyAr 17
1 ATGGCGCGCCGCCAAGCGGTGAAGCGGAGTATGCCAAGCTCGGGCG 50
17 gAlaSerCysLysSerCysArgSerProIleAlaLysAspGlnLeuArgL 34
51 GGCGCTCGTCAAGTATCGCGGTCCCTATCCCAAGGACAGCTCCGTC 100
34 euGlyLysMetValGlnAlaSerGlnPheAspGlyPheMetProMetTrp 50
101 TTGGCAAGATGGTTACAGCGCTCACAGTTCCAGCGGTTCATGCCGATGG 150
51 AsnHisAlaArgCysIlePheSerLysLysAsnGlnIleLysSerValAs 67
151 AACCATGCCAGGTGCATCTTCACAGAGAACACCATGATAAATCCGTGA 200
67 pAspValGluGlyIleAspAlaLeuArgTrpAspGlnGluLysIleA 84
201 CGATGTTGAAGGATAGATGACCTTATAGATGGGATGATCAAGAGAATAC 250
84 rGAsnTyrValGlySerAlaSerAlaGlyThrSerThrAlaAlaPro 100
251 GAAACTACGTTGGGAGTGCCTCAGCTGGTACAAGTTCTACAGCTGCTCT 300
101 ProGluLysCysThrIleGluIleAlaProSerAlaArgThrSerCysAr 117
301 CCTGAGAAATGTACAATTGAGATTGCTCCATCTGCCCGTACTTCATGTAG 350
117 gArgCysSerGluLysIleThrLysGlySerValArgLeuSerAlaLysL 134
351 ACCGATGAGTCAAAAGATTACAAAGGATCGGTCCTCTTCAGCTAAGC 400
134 euGluSerGluGlyProLysGlyIleProTyrPheHisAlaAsnCysPhe 150
401 TTGAGACTGAGGTCCTCCCAAGGATATACCATGGTATCATGCCAACTGTT 450
151 PheGluValSerProSerAlaThrValGluLysPheSerClyTrpAspTh 167
451 TTTGAGGTATCCCGCTGCAACTGTGTGAGAAGTTCTCAGCGTGGGATAC 500
167 rLeuSerAspGluAspLysArgThrMetLeuAspLeuValLysLysAspV 184
501 TTTGTCGATGAGGTCCTCCCAAGGATATACCATGGTATCATGCCAACTGTT 550
184 alGlyAsnAsnGlnAlaAsnLysGlySerLysArgLysLysSerGluAsn 200
551 TTGGCAACATGAACAAATAAGGTTCCCAAGCGCAAGAAAGTGAAT 600

201 AspIleAspSerTyrLysSerAlaArgLeuAspGluSerThrSerGluG 217
601 GATATTGATAGCTACAAATCCGCCAGGTAGATGAAAGTACATCTGAAG 650
217 yThrValArgAsnLysGlyClnLeuValAspProArgGlySerAsnThrS 234
651 TACAGTGCAGAAACAAAGGCGCACTTGTAGACCCACGCTGGTTCATAC 700
234 erSerAlaAspIleGlnLeuLysLeuLysGlnSerAspThrLeuTrp 250
701 GTTCAGCTGATATCCCAACTAAAGCTTAAGGAGCAAAAGTGACACACT 750
251 LysLeuLysAspGlyLeuLysThrHisValSerAlaAlaGluLeuArg 267
751 AAGTTAAAGGATGGACTTAAGACTCATGTATCGGCTGCTGAATTAAG 800
267 pMetLeuGluAlaAsnGlyGlnAspThrSerGlyProGluArgHisLe 284
801 TATGCTTTGAGGCTAATGGGAGGATACATCAGGACCCAGAAAGGCAC 850
284 euAspArgCysAlaAspGlyMetLeuPheGlyAlaLeuGlyProCysP 300
851 TGGATCGCTGTGCGGATGGATGCTATTGTGAGCGCTGGCTTGCCTCC 900
301 ValCysAlaAsnGlyMetTyrTyrAsnGlyClnTyrGlnCysSerG 317
901 GTCTGTGCTAATGGCATGTACTATTATAATGGTCAGTACCAATGCCA 950
317 yAsnValSerGluTrpSerLysCysThrTyrSerAlaThrGluProVal 334
951 TAATGTGTGAGAGTGGTCCAAAGTGTACATCTCGCCACAGAACCTGT 1000
334 rGValLysLysLysTrpGlnIleProHisGlyThrLysAsnAspTyrLe 350
1001 CGGTTAAGCAAGTGGCAATTCACATGCAACAAAGAAATGATTACCT 1050
351 MetLysTrpPheLysSerGlnLysValLysLysProGluArgValLeu 367
1051 ATGAAGTGGTTCAAAATCTCAAAAGGTTAAGAAACCCAGAGGGGTCT 1100
367 oProMetSerProGluLysSerGlySerLysAlaThrGlnArgThrSer 384
1101 ACCAATGTCACTCAGAAATCTGGAAGTAAAGCAACTCAGAGAACATC 1150
384 euLeuSerSerLysGlyLeuAspLysLeuArgPheSerValValClyG 400
1151 TGCTGCTCTCTAAAGGGTGGATAAAATTAAGGTTTCTGTGTAGGACA 1200
401 SerLysGluAlaAlaAsnGluTrpIleGluLysLeuLysLeuAlaGly 417
1201 TCAAAAGAGCAGCAAAATGAGTGGATGAGAAGCTCAAACTTGTGTGC 1250
417 aAsnPheTyrAlaArgValValLysAspIleAspCysLeuIleAlaCys 434
1251 CAACCTCTATGCCAGGGTGTCAAAAGATATTGATTGTTTAAATGCAT 1300
434 lyGluLeuAspAsnGluAsnAlaGluValArgLysAlaArgLeuLys 450
1301 GTGAGCTCGACAATGAAATGCTGAAGTCAGGAAGCAAGGAGGCTGAAG 1350
451 IleProIleValArgGluGlyTyrIleGlyGluCysValLysArgThr 467
1351 ATACCAATGTAAAGGAGGGTTACATTGGAGAAATGTTTAAATAAGAA 1400
467 sCysCysHisLeuIleCysIleAsnTrpAsnAlaLeuGluSerLysG 484
1401 ATGCTGCCATTTGATTGTATAACTGGAATGCCCTTAGAGTCCCTCAA 1450
484 ly***ThrValThrValLysValLysGlyArgSerAlaCysSer*** 500
1451 GCMGTACTGTCTACTGTTAAAGTTAAGGCCGAAGTCTGTTTCATYAA 1500
501 Pro***ValCysLysAsnThrAlaHisIlePro***TrpGluLysHi 517


```

1501 CCTCYGTTTCAAGAATACTGCTACATTCCTTRAGRATGGGAAAAAGCA 1550
|||||
517 sileGlnCys***LeuLysHisValLeuThr***His***ValCysThrG 534
|||||
1551 TATACATGCAACCTTAAACATGTTCTGACCTGNCACNAGGTGGYACAG 1600
|||||
534 lyTyTyValLeuGlnlelleGlnAspGlySerGluCysTyr 550
|||||
1601 GCTACTATGTAATCCAGATCATTTCAACAGGATGATGGCTGAGTGCTAC 1650
|||||
551 ValPheArgLysTrpGlyArgValGlySerGluLysIleGlyGlnLy 567
|||||
1651 GTATTTCTGAAGTGGGACCGGTGGGAGTGCAGAAAAATTGGAGGGCAAAA 1700
|||||
567 sLeuGluGluMetSerLysThrGluAlaIleLysGluPheLysArgLeuP 584
|||||
1701 ACTGAGGAGATGCAAAACTGAGGCAATCAAGGAATTCAAAGATTAT 1750
|||||
584 heLeuGluLysThrGlyAsnSerTrpGluAlaTrpGluCysLysThrAsn 600
|||||
1751 TTCTTGAGAAAGACTGGGAACTCATGGGAAGCTTGGGAATCTAAACCAAT 1800
|||||
601 PheArgLysGlnProGlyArgPheTyrProLeuAspValAspTyrGlyVa 617
|||||
1801 TTTGGAAGCAGCGCTGGGAGATTTTACCCACTTGATGTTGATTATGGTGT 1850
|||||
617 lLysLysAlaProLysArgLysAspIleSerGluMetLysSerSerLeuA 634
|||||
1851 TAAGAAAGCACCACAAACGGAACATATACAGTGAATGAAAGATTCCTTG 1900
|||||
634 laProGlnLeuLeuGluMetLysMetLeuPheAsnValGluThrTyr 650
|||||
1901 CTCTCAATTCGTAGAAGCTCATGAAGATGCTTTTCAATGTGGAGACATAT 1950
|||||
651 ArgAlaAlaMetMetGluPheGlu***AsnMetSerGluMetProLeuGl 667
|||||
1951 AGAGCTCATATGATGGAATTTGAAATTAATGTCAGAAATGCCTCTTGG 2000
|||||
667 yLysLeuSer****AsnIleGlu***GlyPheGluAlaLeuThr**** 684
|||||
2001 GAAGCTAAGCAGGAGGAAATATTCAGRAAGATTTGAAGCATTAACTKRG 2050
|||||
684 *****LeuPheGluGlyHisArg***SerSerThrGlyLeu***Glu 700
|||||
2051 TACMGRATTTTATTGAAGGACACCGCTNATCAAGCACCTGGCTTGTTRGAG 2100
|||||
701 LysAla***LeuLeu*****PheSerLeuLeuSerLeuLe 717
|||||
2101 AAAGCTNAATGTTGTYGAGGCMATSSYTTTTCACCTCTATCCCTTCT 2150
|||||
717 uPheIleLeuIleLeuTyrGlyMetArgMetIleSerTyrSerLysAlaL 734
|||||
2151 ATTCACTCATATATTACGGGATGAGGATGATTTTATATTCNAAGGCCA 2200
|||||
734 ysMetLeuGluAlaLeuGlnAspIleGluIleAlaSerLysIleValGly 750
|||||
2201 AAATGCTTGAAGCTCTGCAGGATATTGAAATTCCTTCAAGATAGTTGCG 2250
|||||
751 PheAspSerAspSerAspGluSerLeuAspAspLysTyrMetLysLeuHi 767
|||||
2251 TTCGATAGGACAGTATGAATCTCTTGTATGATGATAAATATATGAAACTTCA 2300
|||||
767 sCysAspIleThrProLeuAlaHisAspSerGluAspTyrLysLeuIleG 784
|||||
2301 CTGTGACATCACCCGCTGGCTCAGCATAGTGAAGATTACAAGTTAATTG 2350
|||||
784 luGlnTyrLeuLeuAsnThrHisAlaProThrHisLysAspTrpSerLeu 800
|||||
2351 AGCAGTATCTCCTCAACACACATGCTCCTACTCACAAAGGACTGGTGGCTG 2400
|||||
801 GluLeuGluGluValPheSerLeuAspArgAspGlyGluLeuAsnLysTy 817
|||||

```

```

2401 GAACCTGGAGGAAGTTTTTTTCACTTGATCGATGGGAACCTTAATAAGTA 2450
|||||
817 rSerArgTyrLysAsnAsnLeuHisAsnLysMetLeuLeuTrpHisGlys 834
|||||
2451 CTCACGATATAAAATAATCTGCATAACAGATGCTATTATATGGCAGGTT 2500
|||||
834 exArgLeuThrAsnPheValGlyIleLeuSerGlnGlyLeuArgIleAla 850
|||||
2501 CAAGGTTCAGCAATTTTGTGGGAATCTTTAGTCAAGGGCTAAGAATTGCA 2550
|||||
851 ProProGluAlaProValThrGlyTyrMetPheGlyLysGlyLeuTyrPh 867
|||||
2551 CTCTCTGAGGCACCTGTTACTGGCTATATGTTCCGCAAGGCCCTCTACTT 2600
|||||
867 eAlaAspLeuValSerLysSerAlaGlnTyrCysTyrValAspArgAsnA 884
|||||
2601 TGCAGATCTAGTAACCAAGAGCGCAATACATCTGTTATGTGGATAGAATA 2650
|||||
884 snProValGlyLeuMetLeuSerGluValAlaLeuGlyAspMetTyr 900
|||||
2651 ATCCTCTAGGTTTGATGCTTCTTCTGAGGTGCTTTAGGACACATGTAT 2700
|||||
901 GluLeuLysLysAlaThrSerMetAspLysProProArgGlyLysHisSe 917
|||||
2701 GAACCTAAAGAAAGCCACGTCCTCATGGACAAACCTCCAAGAGGGAAGCATTC 2750
|||||
917 rThrLysGlyLeuGlyLysThrValProLeuGluSerGluPheValLysT 934
|||||
2751 GACCAAGGATTTAGCAAAACCGTCCACTGGAGTCAGAGTTTGTCAAGT 2800
|||||
934 rpArgAspAspValValValProCysGlyLysProValProSerSerIle 950
|||||
2801 GCAGGGATGATGCTGCTAGTTCCTTCGCGCAAGCGGTGCCATCATCAATT 2850
|||||
951 ArgSerSerGluLeuMetTyrAsnGluTyrIleValTyrAsnThrSerGl 967
|||||
2851 AGGAGCTCTGAACTCATGTACAATGATGATACATCGTCTACAACACATCCCA 2900
|||||
967 nValLysMetGlnPheLeuLysValArgPheHisLysLysArg 982
|||||
2901 GGTGAAGATGCAGTTCTTCTGCAAGGTGCGTTTCCATCACAGAGG 2946
|||||
seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAZ60618
seq_documentation_block:
ID AAZ60618 standard; DNA; 3212 BP.
XX
AC AAZ60618;
XX
DT 16-MAY-2000 (first entry)
XX
DE DNA encoding the poly(ADP-ribose) polymerase ZAP2 protein of Zea mays;
XX
KW ZAP2; poly(ADP-ribose) polymerase; PARP; poly(ADP-ribose) transferase;
KW programmed cell death; apoptosis; growth rate; stress; cold; pathogen;
KW pest; drought; heat; fungi; nematode; seed-shatter; ss.
XX
OS Zea mays.
XX
FH Key Location/Qualifiers
CDS 81..3023
FT /*tag= a
FT /product= "ZAP2 protein"
XX
PN WO200004173-A1.
XX
PD 27-JAN-2000.
XX
PF 12-JUL-1999; 99WO-EP04940.
XX
PR 17-JUL-1998; 98US-0118276.
XX
PA (PLB2 ) PLANT GENETIC SYSTEMS NV.

```

XX Babiychuk E, Kushnir S, De Block M;
 XX PI
 XX WPI; 2000-182436/16.
 XX P-PSDB; AAY68839.
 XX
 XX Modulating cell death, growth and stress resistance in eukaryotes,
 XX specifically plants, used, e.g. to impart fungus or nematode resistance
 XX
 XX Disclosure; Page 103-108; 126pp; English.

XX The present sequence encodes the ZAP2 protein of Zea mays. This protein
 XX is a poly(ADP-ribose) polymerase (PARP) protein (also known as
 XX poly(ADP-ribose) transferase). The PARP enzyme is involved in programmed
 XX cell death or apoptosis, and is a nuclear enzyme. The ZAP2 polynucleotide
 XX sequences can be used for modulation of programmed cell death in
 XX eukaryotic cells. The method is used, specifically in plants, to induce,
 XX or protect against, programmed cell death, depending on the extent to
 XX which PARP activity is reduced. Reducing expression of endogenous ZAP
 XX class PARP only is also used to modulate programmed cell death, to
 XX increase growth rate and to produce plant cells that are more tolerant
 XX of stress (cold, chemical treatments, pathogens, pests, drought, heat,
 XX etc., or during transformation). Particular applications are generation
 XX of plants that are resistant to fungi or nematodes; are male or female
 XX sterile; or have better seed-shatter properties. The methods are also
 XX used to improve growth of transformed plant cells (and derived calli or
 XX complete plants).

XX Sequence 3212 BP; 974 A; 600 C; 815 G; 823 T; 0 other;

alignment_scores:

Quality: 4740.00 Length: 985
 Ratio: 5.000 Gaps: 5
 Percent Similarity: 96.244 Percent Identity: 95.533

alignment_block:

US-09-236-995D-2 x AAZ60618

Align seg 1/1 to: AAZ60618 from: 1 to: 3212

1 MetAlaAProProLysAlaTrpLysAlaGluTyrAlaLysSerGlyAr 17
 81 ATGGCGCGCGCCAAAGCGTGGAAAGCGGAGTATGCCAAGCTCGGGCG 130
 17 gAlaSerCysLysSerCysArgSerProIleAlaLysAspGlnLeuArgL 34
 131 GGCCTCGTGCAGTCAATGCGGTCCTCCCTATCCCAAGGACCGCTCCGTC 180
 34 euGlyLysMetValGlnAlaSerGlnPheAspGlyPheMetProMetTrp 50
 181 TTGGCAAGATGGTTACAGCGGTACAGTTCGACGGCTTCATGCCGATGTGG 230
 51 AsnHisAlaArgCysIlePheSerTysLysAsnGlnIleLysSerValas 67
 231 AACCATGCCAGGTGCATCTTCAGCAAGAAGAACCCAGATAAAATCCGTGA 280
 67 pAspValGluGlyIleAspAlaLeuArgTrpAspAspGlnGluLysIleA 84
 281 CGATGTTGAAGGATAGATGCACCTTAGATGGGATGATCAAGAGAAGATAC 330
 84 rGAsnTyrValGlySerAlaSerAlaGlyThrSerSerThrAlaAlaPro 100
 331 CAACACTACGTTGGAGTGCTCAGCTGGTACAAAGTTCTACAGCTGCTCT 380
 101 ProGluLysCysThrIleGluIleAlaProSerAlaArgThrSerCysAr 117
 381 CCTGAGAAATGTACAATTGAGATTGTCTCATCTGCCCGTACTTCATGTAG 430
 117 gArgCysSerGluLysIleThrLysGlySerValArgLeuSerAlaLysL 134
 431 ACGATGCGTGAAGGATACAAAAGGATCGGTCCTTTTCAGCTAAGC 480

134 euGluSerGluGlyProLysGlyIleProTrpTyrHisAlaAsnCysPhe 150
 481 TTGAGAGTGAAGGTCCCAAGGGTATACCATGGTATCATGCCAACTGTTTC 530
 151 PheGluValSerProSerAlaThrValGluLysPheSerGlyTrpAspTh 167
 531 TTTGAGGTATCCCGCTGTCAACTGTGTGAGAAGTTCTCAGGCTGGGATAC 580
 167 rLeuSerAspGluAspLysArgThrMetLeuAspLeuValLysLysAspV 184
 581 TTTGTCGCGATCAGGATGAAGAACCATGCTCGATCTTGTATAAAGATG 630
 184 alGlyAsnAsnGluGlnAsnLysGlySerLysArgLysLysSerGluAsn 200
 631 TTGGCAACATGAACAAAATAAGGTTCCAAAGGCCAAGAAAGTGAAT 680
 201 AspIleAspSerTyrLysSerAlaArgLeuAspGluSerThrSerGluG 217
 681 GATATTGATAGTACAAATCCGCCAGGTTAGATGAAGTACATCTGAAGG 730
 217 yThrValArgAsnLysGlyGlnLeuValAspProArgGlySerAsnThrS 234
 731 TACAGTCGGAACAAAGGCAACTTTAGACCCACCGGTGTTCCAACTACTA 780
 234 erSerAlaAspIleGlnLeuLysLysGluGlnSerAspThrLeuTrp 250
 781 GTTCAGCTGATATCCAACTAAAGCTTAAGGAGCAAAAGTACACACTTGG 830
 251 LysLeuLysAspGlyLeuLysThrHisValSerAlaAlaGluLeuArgAs 267
 831 AAGTTAAAGGATGGACTTAAGACTCATGTCGGCTGCTGAATTAAGGGA 880
 267 pMetLeuGluAlaAsnGlyGlnAspThrSerGlyProGluArgHisLeuL 284
 881 TATGCTTGAGGCTAATGGGAGGATACATCAGGACCAGAAAGGCACCTAT 930
 284 euAspArgCysAlaAspGlyMetLeuPheGlyAlaLeuGlyProCysPro 300
 931 TGGATCGCTGTGGGATGGAATGATATTGGAGCGCTGGGTCTTGCCCA 980
 301 ValCysAlaAsnGlyMetTyrTyrTyrAsnGlyGlnTyrGlnCysSerG 317
 981 GTCTGTGCTAATGCGATGCTACTATATTAATGGTCAGTACCAATGCACTG 1030
 317 yAsnValSerGluTrpSerLysCysThrTyrSerAlaThrGluProVala 334
 1031 TAATGTGTGAGTGTGTCAGTGTACATCTCTGCCACAGAACCTGTGCC 1080
 334 rgValLysLysTrpGlnIleProHisGlyThrLysAsnAspTyrLeu 350
 1081 GCGTTAAGAGAAGTGGCAAAATCCACATGGACAAAGAAATGATTACCTT 1130
 351 MetLysTrpPheLysSerGlnLysValLysLysProGluArgValLeuPr 367
 1131 ATGAAGTGGTTCAAAATCTCAAAAGGTTAAGAAACAGAGAGGGTCTTCC 1180
 367 oProMetSerProGluLysSerGlySerLysAlaThrGlnArgThrSerL 384
 1181 ACCAATGTCACTCAGAAATCTGGAAGTAAAGCAACTCAGAGAACATCAT 1230
 384 euLeuSerSerLysGlyLeuAspLysLeuArgPheSerValValGlyCln 400
 1231 TGCTGTCTCTTAAAGGGTTGGATAAAATTAAGGTTTCTGTGTAGGACAA 1280
 401 SerLysGluAlaAlaAsnGluTrpIleGluLysLeuLysLeuAlaGlyAl 417
 1281 TCAAAAGAGCAGCAAAATGAGTGGATTGAGAAGCTCAAACTTGTGTGTGC 1330
 417 aAsnPheTyrAlaArgValValLysAspIleAspCysLeuIleAlaCysG 434
 1331 CAACCTCTATCCAGGGTGTGCAAGATATTGATTGTTTAAATTGCATGTG 1380

```

434  lYGlulLeuAspAsnGluAsnAlaGluValAlaArgArgLeuLys 450
1381 GTGAGTCGACAAATGAAATGCTCAAGTCAGGAAGCAAGGAGCTGAAG 1430
451  ileProileValArgGluGlyTyrIleGlyGluCysVal  LysArgThrL 467
1431 ATACCAATTTAGAGGAGGTTTACATTGGAGAATGTCTTAAAGAAGACAA 1480
467  yScysCysHisLeuIleCysIleAsn...TrpAsnAlaLeuGluSerSerL 483
1481 AATGCTGCCATTTGATTGTATAACTAGAGAAATGCTTAGAGTCTCTCAA 1530
483  ySGly***ThrValThrValLysValLysGlyArgSerAlaCysSer*** 499
1531 AAGGCACTACTGCTACTGTTTAAAGTTTAAAGGGCCGAAAGTGTTCATGAG 1580
500  SerPro***ValCysLysAsnThrAlaHisIlePro***TrpGluLys 516
1581 TCCTCTGGTTTG...CAAGATACGTCTCACATTCCTTGA..AGATGGGAAAA 1626
516  sHisIleGlnCys***LeuLysHisValLeuThr***His***ValCys. 532
1627 GCATATACAAATGCAACCTTAAACATGCTCTGACCTGCG...ACTAGGTGG 1673
533  ThrGlyTyrTyrValLeuGlnIleIleGluGlnAspGlySerGluCy 549
1674 AACAGCTACTATCTACTCCAGATTCATTGAAACAGGATGATGGGTCTGAGTG 1723
549  sTyrValPheArgLysTrpGlyArgValGlySerGluLysIleGlyG 566
1724 CTAGGTATTTCGTAAAGTGGGAGCGGTTGGGAGTGAGAAATTTGAGGGC 1773
566  lnLysLeuGluGluMetSerLysThrGluAlaIleLysGluPheLysArg 582
1774 AAAAAGTGGAGGAGATGTCAAAAGCTGAGGCAATCAAGGAATTCAAAAGA 1823
583  LeuPheLeuGluLysThrGlyAsnSerTrpGluAlaTrpGluCysLysTh 599
1824 TTATTTCTTTGAGAAGACTGGAAGCTCATGGGAGCTTGGGAAATGTAAGAA 1873
599  rAsnPheArgLysGlnProGlyArgPheTyrProLeuAspValAspTyrG 616
1874 CAATTTTCGGAAGCAGCTGGAGATTTTACCACCTTGATGTGTGATATG 1923
616  lyValLysLysAlaProLysArgLysAspIleSerGluMetLysSerSer 632
1924 GTGTTAAGAAAGCAACCAAAACGAAAGATATCAGTGAAATGAAAGTCT 1973
633  LeuAlaProGlnLeuLeuGluMetLysMetLeuPheAsnValGluTh 649
1974 CTGTCTCTCAATGCTAGAACTCATGAAGATGCTTTTCAATGTGGAGAC 2023
649  rTyrArgAlaAlaMetMetGluPheGlu***AsnMetSerGluMetProL 666
2024 ATATAGAGCTGCTATGATGAATTTGAAATTAATATGTCAGAAATGCCTC 2073
666  euGlyLysLeuSer*****AsnIleGlu***GlyPheGluAlaLeuThr 682
2074 TTGGGAAGCTAAGCAAGCAAAATATATGAGAAAGGATTTGAAAGCATTAAC 2123
683  *****LeuPheGluGlyHisArg***SerSerThrGlyLeu** 699
2124 GAGATACAGAAATTTATT..GAAGGACACCGCTGATCAAGCACTGGCTG..TT 2171
699  *GluLysAla***LeuLeu*****PheSerLeuLeuSerL 716
2172 AGAGAAAGCTTAATTTGCTCGGAGCAATCGCTTTTTCACCTTATATCCC 2221
716  euLeuPheIleLeuIleLeuTyrGlyMetArgMetIleSerTyrSerLys 732
2222 TTCTATTCTCATCTCATATATACGGGATGAGCATCATTTG..ATGATCAA 2270
733  AlaLysMetLeuGluAlaLeuGlnAspIleGluIleAlaSerLysIleVa 749

```

```

|||||
2271 GCGAAATGCTTTGAAGCTCTCGAGGATATTGAATTGCTTCAAGATAGT 2320
749  lGlyPheAspSerAspSerAspGluSerLeuAspAspLysTyrMetLysL 766
2321 TGGCTTCGATAGCGACAGTGAATCTCTTGATGATAAATATATGAAAC 2370
766  euHisCysAspIleThrProLeuAlaHisAspSerGluAspTyrLysLeu 782
2371 TTCACITGTGACATCACCCCGCTGGCTCAGATAGTAGAAGATTACAAAGTTA 2420
783  lLeGluGlnTyrLeuLeuAsnThrHisAlaProThrHisLysAspTyrSe 799
2421 ATTGAGCAGTATCTCTCAACACACATGCTCTACTACAGGACTGGTC 2470
799  rLeuGluLeuGluValPheSerLeuAspArgAspGlyGluLeuAsnL 816
2471 GCTGGAACCTGGAGGAAGTTTTTTCACTTTGATCGAGATGAGAACTTAATA 2520
816  ySTyrSerArgTyrLysAsnAsnLeuHisAsnLysMetLeuLeuTrpHis 832
2521 AGTACTCAAGATATATAAATATCTGCATATAAAGATGCTATTATGGCAC 2570
833  GlySerArgLeuThrAsnPheValGlyIleLeuSerGlnGlyLeuArgIl 849
2571 GGTTCAGGTTGACGAATTTTGTGGGAATTTCTAGTCAAGGCTAAGAAT 2620
849  eAlaProProGluAlaProValThrGlyTyrMetPheGlyLysGlyLeuT 866
2621 TGCACCTCTGAGGCACCTGTTACTGGCTATATGTTGCGCAAGGCTCT 2670
866  yrPheAlaAspLeuValSerLysSerAlaGlnTyrCysTyrValAspArg 882
2671 ACTTGCAGATCTAGTAGCAAGAGCGGCAACTACTGTTATGTGGTAGG 2720
883  AsnAsnProValGlyLeuMetLeuLeuSerGluValAlaLeuGlyAspMe 899
2721 AATAATCTGTAGGTTTGATGCTCTTCTGAGGTGCTTTAGGAGACAT 2770
899  tTyrLLeuLysLysAlaThrSerMetAspLysProProArgGlyLysH 916
2771 GTATGAACCTAAAGAAAGCCAGCTCCATGGACAAACCTCCAAGAGGGAAGC 2820
916  isSerThrLysGlyLeuGlyLysThrValProLeuGluSerGluPheVal 932
2821 ATTCGACCAAGGATTAGGCAAAACCTGCCACTGGAGTCAGAGTTTGTG 2870
933  LysTrpArgAspValValProCysGlyLysProValProSerSe 949
2871 AAGTGGAGGATGATGCTGTAGTTCCCTCGCGCAAGCGGTGCCATCATC 2920
949  rIleArgSerSerGluLeuMetTyrAsnGluTyrIleValTyrAsnThrS 966
2921 AATTAGAGCTCTCGAACTCATGTAACATGATGATGATGCTCTACAACACAT 2970
966  erGlnValLysMetGlnPheLeuLeuLysValArgPheHisHisLysArg 982
2971 CCCAGGTGAAGATCAGTCTTCTGCTGAAGGTGCGTTTCCATCACAAGAGG 3020

```

seq_name: /SDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ60615

seq_documentation_block:

ID AAZ60615 standard; DNA; 3211 BP.

XX AAZ60615;

XX 16-MAY-2000 (first entry)

DE DNA encoding the poly(ADP-ribose) polymerase ZAP1 protein of Zea mays.
 KW ZAP1; poly(ADP-ribose) polymerase; PARP; poly(ADP-ribose) transferase;
 KW programmed cell death; apoptosis; growth rate; stress; cold; pathogen;
 KW pest; drought; heat; fungi; nematode; seed-shatter; ss.

XX OS Zea mays.
XX key Location/Qualifiers
XX FH 113..3022
XX CDS
XX FT /*tag= a
XX FT /product= "ZAP1 protein"
XX PN W020004173-Al.
XX PD 27-JAN-2000.
XX PF 12-JUL-1999; 99WO-EP04940.
XX PR 17-JUL-1998; 98US-0118276.
XX PA (PLBZ) PLANT GENETIC SYSTEMS NV.
XX PI Babychuk E, Kushnir S, De Block M;
XX XPI; 2000-182436/16.
XX DR P-PSDB; AAY68833.
XX PT Modulating cell death, growth and stress resistance in eukaryotes,
XX PT specifically plants, used, e.g. to impart fungus or nematode resistance
XX PT
XX PS Example 1; Page 79-84; 125pp; English.
XX CC The present sequence encodes the ZAP1 protein of Zea mays. This protein
CC is a poly(ADP-ribose) polymerase (PARP) protein (also known as
CC poly(ADP-ribose) transferase). The PARP enzyme is involved in programmed
CC cell death or apoptosis, and is a nuclear enzyme. The ZAP1 polynucleotide
CC sequences can be used for modulation of programmed cell death in
CC eukaryotic cells. The method is used, specifically in plants, to induce,
CC or protect against, programmed cell death, depending on the extent to
CC which PARP activity is reduced. Reducing expression of endogenous ZAP
CC class PARP only is also used to modulate programmed cell death, to
CC increase growth rate and to produce plant cells that are more tolerant
CC of stress (cold, chemical treatments, pathogens, pests, drought, heat,
CC etc., or during transformation). Particular applications are generation
CC of plants that are resistant to fungi or nematodes; are male or female
CC sterile; or have better seed-shatter properties. The methods are also
CC used to improve growth of transformed plant cells (and derived calli or
CC complete plants).
XX SQ Sequence 3211 BP; 968 A; 604 C; 813 G; 826 T; 0 other;

alignment_scores:
Quality: 4666.50 Length: 985
Ratio: 4.980 Gaps: 6
Percent Similarity: 95.127 Percent Identity: 94.416

alignment_block:
us-09-236-995d-2 x AAZ60615 ..

Align seg 1/1 to: AAZ60615 from: 1 to: 3211

1 MetAlaLaProProLysAlaTrpLysAlaGluTyrAlaLysSerGlyAr 17
113 ATGGCGGCGCGCCAAAGGCGTGGAGGCGAGTATGCCAAGTCTGGGCG 162

17 qAlaSerCysLysSerCysArgSerProIleAlaLysAspGlnLeuArgL 34
163 GGCCCTCGTGCAGATCATGCCGGTCCCTATCGCCCAAGGACCGCTCCGTC 212

34 euGlyLysMetValGlnAlaSerGlnPheAspGlyPheMetProMetTrp 50
213 TTGGCAAGATGGTTTCAGCGCTCACAGTTCAGCGGCTTCATGCCGATGTGG 262

51 AsnHisAlaArgCysIlePheSerLysLysAsnGlnIleLysSerValas 67
|||||

263 AACCATGCC.....AGCGTTGA 279
67 pAspValGluGlyIleAspAlaLeuArgTrpAspAspGlnGluLysIleA 84
|||||
280 CGATGTTGAAGGGATAGATGCACCTTAGATGGGATGATCAAGAGAGATAC 329
|||||
84 rgAsnTyrValGlySerAlaSerAlaGlyThrSerSerThrAlaAlaPro 100
|||||
330 GAAACACGTTGGAGTGCCTCAGCTGGTACAAGTTCTACAGCTGCTCCT 379
|||||
101 ProGluLysCysThrIleGluIleAlaProSerAlaArgThrSerCysAr 117
|||||
380 CCTGAGAAATGTACAATTTGAGATTGCTTCCATCTGCCCGTACTTTCATGTAG 429
|||||
117 gArqCysSerGluLysIleThrLysGlySerValArgLeuSerAlaLysL 134
|||||
430 ACGATGCGTGAAGAGATTTACAAAAGGATCGGTCCTCTTTTCAGCTAAGC 479
|||||
134 euGluSerGluGlyProLysGlyIleProTrpTyrHisAlaAsnCysPhe 150
|||||
480 TTGAGAGTGAAGGTCCCAAGGTATACCATGGTATCATGCCAACTGTTC 529
|||||
151 PheGluValSerProSerAlaThrValGluLysPheSerGlyTrpAspTh 167
|||||
530 TTTGAGGTATCCCGCTCTGCAACTTTGAGAAGTTCTCAGCGCTGGGATAC 579
|||||
167 rLeuSerAspGluAspLysArgThrMetLeuAspLeuValLysLysAspV 184
|||||
580 TTTGTCGATGAGATTAAGAGACCATGCTCGATCTTTGTTAAAAAAGATG 629
|||||
184 alGlyAsnAsnGluGlnAsnLysGlySerLysArgLysLysSerGluAsp 200
|||||
630 TTGGCAACAATGAACAAATAAGGGTTTCCAAGGCCAAGAAAGTGAATAAT 679
|||||
201 AspIleAspSerTyrLysSerAlaArgLeuAspGluSerThrSerGluGl 217
|||||
680 GATATTGATAGTACAAATTCGCCAGGTAGATGAAGTACATCTGAAGG 729
|||||
217 YThrValArgAsnLysGlyGlnLeuValAspProArgGlySerAsnThrS 234
|||||
730 TACAGTGCAGAAACAAAGGGCAACTTTGAGACCCACGTTGTTCAATACTA 779
|||||
234 erSerAlaAspIleGlnLeuLysLeuLysGluGlnSerAspThrLeuTrp 250
|||||
780 GTTCAGCTGATATCCAACTAAAGCTTAAGGAGCAAAAGTGACACACTTTGG 829
|||||
251 LysLeuLysAspGlyLeuLysThrHisValSerAlaAlaGluLeuArgAs 267
|||||
830 AAGTTAAAGGATGGACTTAAGACTCATGTATCGGCTGCTGAATTAAGGGA 879
|||||
267 pMetLeuGluAlaAsnGlyGlnAspThrSerGlyProGluArgHisLeuL 284
|||||
880 TATGCTTGAGGCTAATGGGAGGATATCATAGGACCCAGAGAGGACACCTAT 929
|||||
284 euAspArgCysAlaAspGlyMetLeuPheGlyAlaLeuGlyProCysPro 300
|||||
930 TGGATCGCTGCGGATGGAATGATATTTGGAGCGCTGGGTCTCTTGCCCA 979
|||||
301 ValCysAlaAsnGlyMetTyrTyrTyrAsnGlyGlnTyrGlnCysSerGl 317
|||||
980 GTCTGCTGCTAATGGCATCTACTATTATTAATGTCAGTACCAATGCAGTGG 1029
|||||
317 YasnValSerGluTrpSerLysCysThrTyrSerAlaThrGluProVala 334
|||||
1030 TAAATGTGCAGAGTGGTCCAAAGTGTACATACTCTGCCACAGAAAGCTGTCC 1079
|||||
334 qValLysLysLysTrpGlnIleProHisGlyThrLysAsnAspTyrLeu 350
|||||
1080 GCGTTAAGAAGAGTGGCAATTCACATGGAACAAGAATGATTACCTT 1129
|||||
351 MetLysTrpPheLysSerGlnLysValLysLysProGluArgValLeuPr 367
|||||
1130 ATGAAGTGTTCAAATCTCAAAAGGTTAAGAAACCCAGAGGGTCTCTTCC 1179
|||||

367 oProMetSerProGluLysSerGlySerLysAlaThrGlnArgThrSerL 384
|||||
1180 ACCAATGTCACCTCAGAAATCTGCAAGTAAAGCAACTCAGAGAACATCAT 1229
384 euLeuSerSerLysGlyLeuAspLysLeuArgPheSerValValGlyGln 400
|||||
1230 TGGTCTCTCTAAAGGGTTGGATAAATTAAGGTTTCTGTGTAGACAA 1279
401 SerLysGluAlaAlaAsnGluTrpIleGluLysLeuLysLeuAlaGlyAl 417
|||||
1280 TCAAAAGACAGCAATGATGATGGATGAGAGCTCAAACTTCTGGTGC 1329
417 aAsnPheTyrAlaArgValValLysAspLysAspCysLeuIleAlaCysG 434
|||||
1330 CAACCTCTATGCCAGGGTTGTCAAGATATATGTTTAAATGTCATGTG 1379
434 lyGluLeuAspAsnGluAsnAlaGluValArgLysAlaArgLeuLys 450
|||||
1380 GTGAGCTCGCAATGAATAATGCTGAAGTCAGGAAGCAAGGAGGCTGAAG 1429
451 IleProIleValArgGluGlyTyrIleGlyGluCysVal. LysArgThrL 467
|||||
1430 ATACCAATGCTAAGGGAGGGTTACATGAGAGATGTGTTAAAGAACAA 1479
467 yScysCysHisLeuIleCysIleAsn...TrpAsnAlaLeuGluSerSerL 483
|||||
1480 AATGCTGCCATTTGATTTGTATAACTAGAGATGCTTAGAGTCTCTCAA 1529
483 yGlyLysThrValThrValLysValLysGlyArgSerAlaCysSer*** 499
|||||
1530 AAGCGAGTACTGTCACCTGTAAAGTTAAGGGCCGAGGCTGCTGTCATGAG 1579
500 SerPro***ValCysLysAsnThrAlaHisIlePro***TrpGluLys 516
||| : : : : : |||||
1580 TCCTCTGGTTG...CAAGATAGCTCCTACATTCITGA. AGATGGGAAA 1625
516 sHisIleGlnCys***LeuLysHisValLeuThr***His***ValCys. 532
|||||
1626 GCATATACAACTGCACTTAAACATGTCTGACCTGGC...ACTAGTGTG 1672
533 ThrGlyTyrTyrValLeuGlnIleIleGluGlnAspGlySerGluCys 549
: : : : : |||||
1673 AACAGCTACTATGTACTCCAGATCATTTGAACAGAGATGATGGTCTGAGTG 1722
549 sTyrValPheArgLysTrpGlyArgValGlySerGluLysIleGlyGlyG 566
|||||
1723 CTACGTATTTCTGTAAGTGGGGACGGGTTGGGAGTGAGAAAATTCGAGGCG 1772
566 loLysLeuGluGluMetSerLysThrGluAlaIleLysGluPheLysArg 582
|||||
1773 AAAAAGTGGAGAGATGTCAAAACTGAGGCAATCAAGGAATTCAAAAGA 1822
583 LeuPheLeuGluLysThrGlyAsnSerTrpGluAlaTrpGluCysLysTh 599
|||||
1823 TTATTTCTTGAGAGACTTGAAGACTCATGGGAAGCTTGGGAATGTAAAC 1872
599 rAsnPheArgLysGlnProGlyArgPheTyrProLeuAspValAspTyrG 616
|||||
1873 CAATTTTCGGAAGCAGCTGGGAGATTTTACCCACTTGTATGTTGATTATG 1922
616 lyValLysLysAlaProLysArgLysAspIleSerGluMetLysSerSer 632
|||||
1923 GTGTTAAGAAAGCACCACCAAGAAAGATATACAGTGAATGAAAGTTCT 1972
633 LeuAlaProGlnLeuLeuGluLeuMetLysMetLeuPheAsnValGluTh 649
|||||
1973 CTTGCTCTCAATTCCTAGAACTCATGAAGATGCTTTTCAATGTGGAGAC 2022
649 rTyrArgAlaAlaMetMetGluPheGlu***AsnMetSerGluMetProL 666
|||||
2023 ATATAGAGCTGCTATGATGGAATTTGAAATTAATATGTAGAAATGCTC 2072

666 euGlyLysLeuSerSer*****AsnIleGlu***GlyPheGluAlaLeuThr 682
|||||
2073 TTGGGAGCTAAGCAAGGAATAATTTGAGAAAGGATTTGAGACATTAAC 2122
683 *****LeuPheGluGlyHisArg***SerSerThrGlyLeu** 699
||| : : : : : |||||
2123 GAGATACAGAAATTTATT. GAAGGACACCGCTGATCAAGCACTGGCTG. TT 2170
699 *GluLysAla***LeuLeu*****PheSerLeuLeuSerL 716
|||||
2171 AGAGAAAGCTTAATTTGCTGCGAGCAATCGCTTTTCACTCTTATCCC 2220
716 euLeuPheIleLeuIleLeuTyrGlyMetArgMetIleSerTyrSerLys 732
|||||
2221 TTCTATTCATCCTCATATTTATACGGGATGAGGATGATTG. ATGATCAA 2269
733 AlaLysMetLeuGluAlaLeuGlnAspIleGluIleAlaSerLysIleVa 749
|||||
2270 GCGAAATGCTTGAAGCTCTGCAGGATATTGAAATTTGCTTCAAGATAGT 2319
749 lGlyPheAspSerAspSerAspGluSerLeuAspLysTyrMetLysL 766
|||||
2320 TGGCTTCGATACGACAGCTGATGAATCTCTTGATGATAAATATATGAAC 2369
766 euHisCysAspIleThrProLeuAlaHisAspSerGluAspTyrLysLeu 782
|||||
2370 TTCCTGTGACATCACCCCGCTGCTCAGCATAGTGAAGATTACAAGTTA 2419
783 IleGluGlnTyrLeuLeuAsnThrHisAlaProThrHisLysAspTrpSe 799
|||||
2420 ATTGACAGTATCTCTCAACACACATGCTCTACTCACAAGGACTGGTC 2469
799 rLeuGluLeuGluGluValPheSerLeuAspArgAspGlyGluLeuAsnL 816
|||||
2470 CTTGGAAGCTGGAGGAAGTTTTTTCATTCGATCGAGATGGAGAACTTAATA 2519
816 yTyrSerArgTyrLysAsnAsnLeuHisAsnLysMetLeuLeuTrpHis 832
|||||
2520 AGTACTCAAGATATAAAAAATAATCTGCATAACAAGATGCTATTATGGCAC 2569
833 GlySerArgLeuThrAsnPheValGlyIleLeuSerGlnGlyLeuArgII 849
|||||
2570 GGTTCAAGTTGACGAATTTTGTGGGAATTTCTTAGTCAAGGGCTAAGAAT 2619
849 eAlaProProGluAlaProValThrGlyTyrMetPheGlyLysGlyLeu 866
|||||
2620 TGCACCTCTCTGAGGCACCTGTTACTGCTATATGTTGCGCAAGGCCCTCT 2669
866 yrPheAlaAspLeuValSerLysSerAlaGlnTyrCystyrValAspArg 882
|||||
2670 ACTTGCAGATCTAGTAAGCAAGAGCGCACAACTACTGTTATGTGATAGG 2719
883 AsnAsnProValGlyLeuMetLeuLeuSerGluValAlaLeuGlyAspMe 899
|||||
2720 AATACTCTGTAGTTTGTATGCTTCTTCTGAGGTTGCTTTTAGGAGACAT 2769
899 tTyrGluLeuLysLysAlaThrSerMetAspLysProProArgGlyLysH 916
|||||
2770 GTATGAACATAAAGAAAGCCAGCTCCATGGACAAACCTCCAAAGAGGAAGC 2819
916 isSerThrLysGlyLeuGlyLysThrValProLeuGluSerGluPheVal 932
|||||
2820 ATTGACCAAGGGATTAGGCAAAACCGCTGCACTGGAGTCAGAGTTTGTG 2869
933 LysTrpArgAspValValValProCysGlyLysProValProSerSe 949
|||||
2870 AGTGGAGGATGATGCTGCTAGTTCCCTGGGCAAGCCGCTGCATCATC 2919
949 rIleArgSerSerGluLeuMetTyrAsnGluTyrIleValTyrAsnThrS 966
|||||
2920 AATTAGGAGCTCTGAATCATCATGATGATGATGATGATGATGATGATGAT 2969
966 erGlnValLysMetGlnPheLeuLysValArgPheHisLysArg 982

```
|||||
2970 CCCAGGTGAAGTGCAGTCTTGTGCTGAAGGTGCGTTTCCATCAAGAGG 3019
seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1996.DAT:AAT13732
seq_documentation_block:
ID AAT13732 standard; DNA; 3793 BP.
XX
AC AAT13732;
XX
XX 31-OCT-1996 (first entry)
XX
DE DNA-binding domain of poly(ADP-ribose) polymerase coding sequence.
XX
XX PARG; DNA-binding domain; poly(ADP-ribose) polymerase; gene therapy;
XX
KW tumour treatment; DNA repair; over-expression; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
CDS 96..3137
FT /*tag= a
FT /product= poly(ADP-ribose)_polymerase
FT /note= "PARG"
FT misc_feature 67..1220
FT /*tag= b
FT /note= "encodes DNA-binding domain"
XX
PN W09618737-A2.
XX
XX 20-JUN-1996.
XX
XX 15-DEC-1995; 95WO-DE01817.
XX
XX 16-DEC-1994; 94DE-4444949.
XX
XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
XX
XX Buerkle A, Kuepper J, Zur Hausen H;
XX
XX WPI; 1996-300654/30.
XX
XX P-PSDB; AAR99642.
XX
XX Vectors contg. insert encoding DNA-binding domain of
XX
XX poly(ADP-ribose) polymerase - useful for gene therapy, esp. of
XX
XX tumours
XX
XX Claim 4; Fig 1; 22pp; German.
XX
XX The present sequence encodes a poly(ADP-ribose) polymerase (PARG) contg.
XX
XX a DNA-binding domain (DBD). The DNA fragment from -29 to + 1127
XX
XX (nucleotides 67-1220 of this sequence) encoding the DBD can be inserted
XX
XX into vectors which are used for gene therapy. Over-expression of the DBD
XX
XX inhibits the DNA repair function of PARG, so the vectors are useful for
XX
XX gene therapy or tumours, esp. in combination with conventional chemo-
XX
XX and/or radiotherapy.
XX
XX Sequence 3793 BP; 1049 A; 847 C; 1034 G; 863 T; 0 other;
XX
SQ
alignment_scores:
Quality: 1594.50 Length: 1070
Ratio: 2.311 Gaps: 37
Percent Similarity: 64.486 Percent Identity: 36.355
alignment_block:
US-09-236-995D-2 x AAT13732 ..
Align seg 1/1 to: AAT13732 from: 1 to: 3793
6 LysAlaTrpLysAlaGluTyrAlaLysSerGlyArgAlaSerCysLysSe 22
|||||
114 AAGCTCTATCGAGTCGAGTAGCCGACGAGCGCGCCCTCTTGCAAGAA 163
```

```
22 rCysArgSerProIleAlaLysAspGlnLeuArgLeuGlyLysMetValc 39
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
164 ATGCGCAGAGCATCCCCAAGACTCGCTCCGATGGCCATCATGTGTC 213
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
39 InAlaSerGlnPheAspGlyPheMetProMetTrpAsnHisAlaArgCys 55
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
214 AGTCGCCCATGTTGTATGTAAGAAAGTCCACACTGGTACCACCTCTCTGTC 263
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
56 IlePheSerLysLysAsnGlnIleLysSerValAsp... AspValGluG1 71
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
264 TTCTGGAAGTGGGCCACTCCATCCGGCACCCCTGACGTTGAGGTGGATGC 313
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
71 YIleAspAlaLeuArgTrpAspGlnGluLysIleArgAsnTyrVal. 87
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
314 GTTCTCTGAGTTCGGTGGATGACACAGCAAGAAAGTCAAGAAGACGCG 363
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
88 .....GlySerAlaSerAlaGlyThrSerSerThrAlaAlaProPro 101
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
364 AAGCTGGAGGAGTGACAGGCAAGGCCAGGATGCAATTGGTCAAGGCA 413
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
102 GluLysCys.....ThrIleGluIleAlaProSerAlaArgTh 114
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
414 GAGAAGACTCTGGGTGACTTTCAGCAGAGATATGCCAAGTCCACACAGAAG 463
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
114 rSerCysArgArgCysSerGluLysIleThrLysGlySerValArgLeuS 131
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
464 TACGTGCAAGGGGTGTATGGAGAAGATAGAAAGGGCCAGGTGCGCTGT 513
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
131 erAlaLysLeu...GluSerGluGlyProLys...GlyIle.....Pro 143
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
514 CCAAGAGAGTGTGGACCCGAGGAAAGCCACAGCTAGGATGATGACCGC 563
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
144 TrpTyrHisAlaAsnCysPhePheGlu.....Valse 154
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
564 TGTACCATCCAGCTGCTTCTCAAGAACAGGAGGAGCTGGGTTTCCG 613
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
154 rProSerAlaThrValGluLysPheSerGlyTrpAspThrLeuSerAspG 171
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
614 GCCCGAGTACAGTGCAGTCAAGTCAAGGGCTTCAGCTCCTTCTGCTACAG 663
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
171 luAspLysArgThrMetLeuAspLeuValLysLysAsp.....Val 184
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
664 AGGATAAA.....GAAGCCCTGAAGAACAGCAGCTCCACGAGGATC 701
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
185 GlyAsnAsnGluGlnAsnLysGlySerLys..... 194
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
702 AAGAGTGAAGGAAGAGAAAGGGGATAAGCTGGATGGAGTGAAGT 751
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
195 ....ArgLysLysSerGluAsnAspIleAspSerTyrLysSerAlaArgL 210
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
752 GCGCAAGAAGAAATCTAAAAAAGAAAGAAC.....AAGGATAGTAAGC 795
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
210 euAspGluSerThrSerGluGlyThrValArgAsnLysGlyGlnLeuVal 226
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
796 TTGAAGAA..... 803
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
227 AspProArgGlySerAsnThrSerSerAlaAspIleGlnLeuLysLeuLy 243
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
804 .....CGCTAAA 811
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
243 sGluGlnSerAspThrLeuTrpLysLysAspGlyLeuLysThrHisV 260
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
812 GGCTCAGAACGACCTGATCTGGAAACATCAAGGACGAGCTAAAGAAGTGT 861
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
260 alSerAlaAlaGluLeuArgAspMetLeuGluAlaAsnGlyGlnAspThr 276
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
862 GTTCAACTAATGACCTGAAGGAGGTACTCTTCAACAAAGCAGCAAGTG 911
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
277 SerGlyProGluArgHisLeuLeuAspArgCysAlaAspGlyMetLeuPh 293
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
912 CCTTCTGGGAGTGGCGGATCTTGGACCGGAGTAGCTGATGGCATGTGT 961
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```


2666 ATTGCTGGGACGGGTCCAGGACCACTTGTGCTGGGATCTCTGCTCC 2715
 845 InGlyLeuArgIleAlaProProGluAlaProValThrGlyTyrMetPhe 861
 2716 AGGCTTCGGATAGCCCGCTGAAGCGCGGTGACAGGTACATGTTT 2765
 862 GlyLysGlyLeuTyrPheAlaAspLeuValSerLysSerAlaGlnTyrCy 878
 2766 GCTAAGGAGTCTATTTCGCTGACATGCTCCAAAGAGTGCCAACTACTG 2815
 878 sTyrValAspArgAsnAsnProValGlyLeuMetLeuLeuSerGluVala 895
 2816 CCATACGCTCTCAGGAGACCCCAATAGGCTTAATCCTGTGGGAGAAGTTG 2865
 895 IaLeuGlyAspMetTyrGluLeuLysLysAlaThrSerMetAspLysPro 911
 2866 CCCTTGGAAACATGATGACTGAAGCAGCTTTCACATATCAGCAAGTTA 2915
 912 ProArgGlyLysHisSerThrLysGlyLeuGlyLysThrValProLeuG1 928
 2916 CCCAAGGCGACAGCAGTGTCAAGGTTTGGCAAACTACCCCTGATCC 2965
 928 uSerGluPheValLysTyrArgAspAspValValProCysGlyLysP 945
 2966 TTCACCTAACATTAAGT...CTGGATGGGTAGACGTTCTCTTGGGACCG 3012
 945 roValProSerSerIleArgSerSerGluLeuMetTyrAsnGluTyrIle 961
 3013 GGATTTTCATCTGGGTGTGAATGACACCTCTCTACTATATAACGAGTACAT 3062
 962 ValTyrAsnThrSerGlnValLysMetGlnPheLeuLeuLysValArgPh 978
 3063 GTCATGATATGCTCAGGTAAATCTGAAGTATCTGCTGAAACTGAAATT 3112
 978 ehIshisLys 981
 3113 CAATTTTAAG 3122

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAF63954

seq_documentation_block:

ID AAF63954 standard; DNA; 3045 BP.

XX AAF63954;

XX 05-APR-2001 (first entry)

DE Human tankyrase2 related coding sequence SEQ ID NO: 136.

XX Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasia; aging;
 KW inflammatory disorder; ds.

XX Homo sapiens.

XX WO200100849-A1.

XX 04-JAN-2001.

XX 28-JUN-2000; 2000WO-US17827.

XX 29-JUN-1999; 99US-0141582.

XX (ICOS-) ICOS CORP.

XX Christenson E, Demaggio AJ, Goldman PS, McElligott DL;

XX WPI; 2001-102896/11.

DR P-PSDB; AAB66296.

XX New tankyrase2 polypeptides, useful for treating conditions mediated by
 PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
 PT inflammatory and autoimmune disorders -

XX

PS Example 2; Page 203-207; 242pp; English.

XX The present invention provides the protein and coding sequence for the
 CC human tankyrase2 protein. This is found in two different versions,
 CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has
 CC polyADP-ribosylation activity and is involved in the modification of
 CC TRF1, which is a telomere-specific binding protein. The regulation of
 CC telomere length, in which TRF1 has a role, is linked to ageing and
 CC cancer. The sequences are useful in the treatment of cancers and
 CC inflammatory disorders.

XX Sequence 3045 BP; 863 A; 710 C; 851 G; 621 T; 0 other;

alignment_scores:

Quality: 1584.00 Length: 1060

Ratio: 2.289 Gaps: 36

Percent Similarity: 65.283 Percent Identity: 36.226

alignment_block:

US-09-236-995D-2 x AAF63954 ..

Align seg 1/1 to: AAF63954 from: 1 to: 3045

6 LysAlaTrpLysAlaGluTyrAlaLysSerGlyArgAlaSerCysLysSe 22
 19 AAGCTCTATCGAGTCGAGTACGCCAAGAGCGGCGCGCTCTTGCAGAA 68
 22 rCysArgSerProIleAlaLysAspGlnLeuArgLeuGlyLysMetValG 39
 69 ATCAGCAGAGCATCCCAAGGACTCGCTCCGGATGGCCATCATGGTGC 118
 39 InAlaSerGlnPheAspGlyPheMetProMetTrpAsnHisAlaArgCys 55
 119 AGTCGCCCATGTTGATGGAAGTCCACACTGGTACCACCTCTCCTCG 168
 56 IlePheSerLysLysAsnGlnLysSerValAsp...AspValGluG1 71
 169 TTCTGGAAGGTGGGCACCTCCATCCGCGCACCTGACGTTGAGTGGATGG 218
 71 YlleAspAlaLeuArgTrpAspAspGlnGluLysIleArgAsnTyrVal. 87
 219 GTTCTCTGAGCTTCGGTGGGATGACCAGCAGAAAGTCAAGAACAGACGCG 268
 88GlySerAlaSerAlaGlyThrSerSerThrAlaAlaProPro 101
 269 AAGCTGGAGGAGTGACAGCAAGGCCAGGATGGAATTGGTAGCAAGGCA 318
 102 GluLysCys.....ThrIleGluIleAlaProSerAlaArgTh 114
 319 GAGAAGACTCTGGGTGACTTTGAGCAGAGATGATGCCAAGTCCCAACAAG 368
 114 rSerCysArgArgCysSerGluLysIleThrLysGlySerValArgLeuS 131
 369 TACGTGCAAGGGGTGTATGGAGAGATAGAAAGGCCAGGCGGCTGT 418
 131 erAlaLysLeu...GluSerGluGlyProLys...GlyIle.....Pro 143
 419 CCAAGAAGATGGTGGACCCGAGAGACAGCTAGGCATGATTGACCGC 468
 144 TrpTyrHisAlaAsnCysPhePheGlu.....ValSe 154
 469 TGTATACCTCCAGGCTGCTTTGTCAAGAACAGGAGGAGCTGGGTTTCG 518
 154 rProSerAlaThrValGluLysPheSerGlyTrpAspThrLeuSerAspG 171
 519 GCCCGAGTACAGTGGCAGCTCAAGGCTCAGGCTTCAGCCTCTCTGTACAG 568
 171 luAspLysArgThrMetLeuAspLeuValLysLysAspValGlyAsnAsn 187
 569 AGCATATAA.....GAAGCCCTGAAGAAGCAGCTCCAGGAGTC 606
 188 GluGlnAsnLysGlySerLysArgLysLysSerGluAsnAspIleAspSe 204

607 AAG.....AGTGAAGAAAGAGAAAGGCGATGAGGTGGATGG 644
204 rTyrLysSerAlaArgLeuAspGluSerThrSerGluGlyThrValArgA 221
645 A.....GTGGATCAAGTGGCGAAGAGAAATCTAAAGAG 679
221 snLysGlyGlnLeuValAspProArgGlySerAsnThrSerSerAlaAsp 237
680 AAAAGACAG.....GATAGTAAG 699
238 IleGlnLeuLysLeuLysGluGlnSerAspThrLeuTrpLysLeuLysAs 254
700 CTTGAAAAAGCCCTAAAGGCTCAGACGACCTGCATCTGGAACATCAAGGA 749
254 pGlyLeuLysThrHisValSerAlaAlaGluLeuArgAspMetLeuGluA 271
750 CGAGCTAAAGAAAGTGTTCACCTAATGACCTGAAGGAGCTACTCATCT 799
271 laAsnGlyGlnAspThrSerGlyProGluArgHisLeuLeuAspArgCys 287
800 TCAACAGCAGCAAGTGCCTCTCGGGAGTCGCGGATCTTGGACCGAGTA 849
288 AlaAspGlyMetLeuPheGlyAlaLeuGlyProCysProValCysAlaAs 304
850 CTTGATGGATGTTGTTGGTCCCTTCCTTCCCTGGAGGAATGCTCGGG 899
304 nGlyMetTyrTyrTyrAsnGlyGlnTyrGlnCysSerGlyAsnValSerG 321
900 TCAGCTGGTCTTCAAGACGCGATGCCTATTACTGCACCTGGGGACCTCAGT 949
321 luTrpSerLysCysThrTyrSerAlaThrGluProValArgValLysLys 337
950 CTTGGACCAAGGTATGGTCAAGACACACACACCAACCGG.....AAG 993
338 LysTrpGlnIleProHisGlyThrLysAsnAspTyrLeuMetLysTrpPh 354
994 GAGTGGTAACCCCA.....AAGGAATCCGAGAAATCTTACTCT 1034
354 eLysSerGlnLysValLysLysProGluArgValLeuProMetSer. 370
1035 CAAGAAATGAAGGTTAAAGACGAGCGATATTATCCCGCCAGAAACCA 1084
371ProGluLysSerGlySerLysAla 378
1085 GCGCTCCGTCGGCGCCAGCGCTCCGCGCTCCACAGCCTCGGCTCCTGCT 1134
379 ThrGlnArgThrSerLeuSerSerLysGlyLeuAspLysLeuArgPh 395
1135 GCTGTGAACCTCTCTGCTTACAGCATAGCCATTATCCACATGAAGAT 1184
395 eSerValValGly.....GlnSerLysGluAlaAlaAsnGluTrpI 409
1185 CTTGACTCTCGGGAAGTGTCCCGGAACAAGGATGAAGTGAAGGCCAIGA 1234
409 leGluLysLeu.....LysLeuAlaGlyAlaAsnPheTyrAlaArgVal 423
1235 TTGAAAATTCGGGGGAGTGTAGCGGGGCGGCGCAACAAGGCTTCCCTG 1284
424 ValLysAspIleAspCysLeuIleAlaCysGlyGluLeuAspAsnGluAs 440
1285TGCAATCAGCCACCAAAAGGAGGTGGAAAGATGAA 1319
440 nAlaGluValArgLysAlaArgLeuLysIleProIleValArgGluG 457
1320 TAAGAAGATGAGGAAGTAAGGAAGCAACATCCGAGTTGTGTGAGG 1369
457 lyTyrlleGlyCysValLysArgThrLys..... 467
1370 ACTTCCTCCAGAGCTCTCCGCTCCACCAAGAGCCTTCAGGATTTGTT 1419
468 CysCysHisLeuIleCysIleAsnTrpAsnAlaLeuGluSerLysG 484
..... 499

1420 TTAGCGCATCTTGTCC...CCTTGGGGGCGAGAGTGAAGCGAGAGCC 1466
484 y***ThrValThrValLysValLysGlyArgSerAlaCysSer***SerP 501
1467 TGTTCGAAGTT...GTGGCCCAAGAGGAAGTCAAGGCGCTGCGTCTCCA 1513
501 to***ValCysLysAsnThrAlaHisIlePro****Trp..... 514
1514 AA.....AAAGCAAGGCCAGGTCAGGAGGAAGGTATCAACAAA 1554
515 ...GluLysHisIleGlnCys***LeuLys..... 523
1555 TCTGAAAAGAGATTAATACTCTTAAAGGAGGAGCAGCTGTGGATCC 1604
524HisValLeu..... 526
1605 TGATTCTGGACTGGACACTCTGGCATGTCTCGAGAAAGGTGGGAGG 1654
527Thr***His***Val.....CysThrGly 534
1655 TCTTCAGTGCACCCCTTGGCCTGGTGGACATCGTTAAAGGAACCAACTCC 1704
535 TyrTyrValLeuGlnIleGluGlnAspAspGlySerGluCysTyrVa 551
1705 TACTACAAAGCTGCACCTTCTGGAGGAGCACAAGGAAAACAGGTATTGGAT 1754
551 lPheArgLysTrpGlyArgValGlySerGluLysIleGlyGlnLysL 568
1755 ATTCAAGTCTGGGGCGCTGTGGGTAGC...GTATCGGTAGCAACAAAC 1801
568 euGluGluMet...SerLysThrGluAlaIleLysGluPheLysArgLeu 583
1802 TGAACACAGATCCCGTCCAAGGAGGATGCCATTGAGCAGTTTCATGAATTA 1851
584 PheLeuGluLysThrGlyAsnSerTrpGluAlaTrpGluCysLysThrAs 600
1852 TATGAAGAAAACCGGGAACGCTTGGCACTCC.....AAAAA 1889
600 nPheArgLysGlnProGlyArgPheTyrProLeuAspValAspTyrGlyV 617
1890 TTTTCAAGNAGTATCCCAAAAAGTTTACCCTCGGAGATTGACTATGGCC 1939
617 al...LysLysAlaProLysArgLysAspLysSerGlu...MetLysSer 631
1940 AGGATGAAGAGGAGTGAAGAGCTACAGTAAATCTTGGCACCACAGTCC 1989
632 SerLeuAlaProGlnLeuLeuGluLeuMetLysMetLeuPheAsnValG 648
1990 AAGCTCCCAAGCCAGTTTCAGGACCTCATCAAGATGATCTTTGATGTGA 2039
648 uThrTyrArgAlaAlaMetMetGluPheGlu***AsnMetSerGluMetP 665
2040 AAGTATCAAGAAAGCCATGTGGAGTATGAGATCGACCTTCAGAAGATGC 2089
665 roLeuGlyLysLeuSer*****AsnIleGlu***GlyPheGluAlaLeu 681
2090 CCTTGGGAAGTCAAGCAAAAAGGAGATCCAGCGCATCTCCATCCCTC 2139
682 Thr*****LeuPheGluGlyHisArg***SerSerThrGlyLe 698
2140 AGT.....GAGGTCCAGCAGCGGTGTCTCAGGGCGAG 2171
698 u***GluLysAla***LeuLeu*****PheSerLeuLeuS 715
2172 CAGGACTCTCAGATCTGTGATCTCTCAATCGCTTTTACACCTGATCC 2221
715 erLeuLeuPheIleLeuIleLeuTyrGlyMetArgMet..... 727
2222 CCCACGACTTT.....GGGATGAAGAGAGCTCCGCTCCTG 2256
728IleSerTyrSerLysAlaLysMetLeuGluAlaLeuG 740
2257 AACAAATGCAGACAGTGTGCAGGCCAAGGTGGAATGCTTGACAAACCTGCT 2306

740 nAspIleGluLeuAlaSerTyrIleVal.....GlyPheAspSerAsps 755
2307 GGACATGAGTGGCTACAGCTCTGCTCAGGGAGGGTCTGATGATGACGA 2356
755 erAspGluSerLeuAspAspIleTyrMetLeuHisCysAspIleThr 771
2357 GCAAGGATCCCATCGATGCTCAAGCTATGAGAGCTCAAAAGTGCATTAA 2406
772 ProLeuAlaHisAspSerGlyLeuHisLeuIleGluGlnTyrLeuLe 788
2407 GTGGTTCACAGATCTGAGAGCCGAGATCATCAGGAAGTATGTAA 2456
788 uAsnThrHisAlaProThrHisLysAspTrpSerLeuGluGluVal 805
2457 GAACACTCATGCAACACACACAGCTGCGTATGACTTGAAGTATCGATA 2506
805 alPheSerLeuAspArgAspGlyGluLeuAsnLysTyrSerArgTyrLys 821
2507 TCTTTAGATAGAGCTGAGCGGATGCTGCGGATAGCCCGCTGAGCGC 2556
822 AsnAsnLeuHisAsnLysMetLeuLeuTrpHisGlySerArgLeuThrAs 838
2557 ...CAGCTTCATAACCGAAGATGCTGTGGCAGCGGTCCAGGACCA 2603
838 nPheValGlyIleLeuSerGlnGlyLeuArgIleAlaProGluAlap 855
2604 CTTTGTGGGATCCCTGTCGAGGGTCTTCCGATAGCCCGCTGAGCGC 2653
855 roValThrGlyTyrMetPheGlyLysGlyLeuTyrPheAlaAspLeuVal 871
2654 CCGTCACAGGCTACATGTTGTAAGGATCTATTTCGCTGACATGTC 2703
872 SerLysSerAlaGlnTyrCysTyrValAspArgAsnProValGlyLe 888
2704 TCCAAGAGTGCCAACTACTACATACGCTCTCAGGGAGACCCCAATAGGCT 2753
888 uMetLeuLeuSerGluValAlaLeuGlyAspMetTyrGluLeuLys 905
2754 AATCCTGTGGGAGAGTTCCTTGGAAACATGTATGAAGTGAAGCAGC 2803
905 laThrSerMetAspLysProProArgGlyLysHisSerThrLysGlyLeu 921
2804 CTTACATATACAGAGTTTACCCAGGGCAAGCACAGTCTCAAGAGTTG 2853
922 GlyLysThrValProLeuGluSerGluPheValLysTrpArgAspVa 938
2854 GGCAAACTACCCCTGATCCTTCAGCTAACATTAGT...CTGGATGGTGT 2900
938 lValValProCysGlyLysProValProSerSerIleArgSerGluL 955
2901 AGACGTTCTCTTGGACCGGATTCATCTGCTGTGATAGACCTCTC 2950
955 euMetTyrAsnGluTyrIleValTyrAsnThrSerGlnValLysMetGln 971
2951 TACTATATAACGAGTACATGTCTATGATATTGCTCAGGTAATCTGAAG 3000
972 PheLeuLeuLysValArgPheHisHisLys 981
3001 TATCTGCTGAACCTGAATTCATTTTAA 3030
seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT:AA223799
seq_documentation_block:
ID AA223799 standard; cDNA: 3580 BP.
XX
AC
XX
XX
DT
XX
DE
XX
KW
Gene therapy; poly(adenosine diphosphate-ribose) polymerase; treatment;

PARP; antitumor; nuclear DNA repair; proliferating cell; DNA damage; protection; genomic instability; cancer; prevention; human; ds.
Homo sapiens.
Key Location/Qualifiers
CDS 140..3184
/*tag= a
/product= "poly(ADP-ribose) polymerase"
DE19808889-A1.
XX
PN 09-SEP-1999.
XX
PD 03-MAR-1998; 98DE-1008889.
XX
PF 03-MAR-1998; 98DE-1008889.
XX
PR (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
XX
PA Buerkle A, Meyer R;
PI WPI: 1999-509563/43.
DR P-PSDB; AAY33699.
XX
XX New gene therapy vector expressing poly(adenosine diphosphate-ribose)-polymerase for treating or preventing tumors -
PT Claim 4; Fig 1; 12pp; German.
XX
CC This invention describes a novel vector (A), suitable for gene therapy, which contains a DNA insert (I) that encodes a practically complete poly(adenosine diphosphate-ribose) polymerase (PARP) which has antitumor activity. PARP is a nuclear DNA repair enzyme (activated by strand breakage), that allows recovery of proliferating cells from the toxic effects of DNA damage (caused by alkylating or oxidizing agents or radiation), and protects cells against such damage or other causes of genomic instability. (A) are used for treatment of cancer and for cancer prevention in subjects at high risk (e.g. those with tumor-associated genetic defects). Tumor cells treated with (A) show an increased tendency to die when treated with radiation or chemotherapeutic agent, and in cells that survive this treatment they inhibit genomic instability, so should reduce the likelihood of further development of both malignant cells and resistance to chemotherapy. This sequence encodes the human poly(ADP-ribose) polymerase described in the invention.
XX
SQ Sequence 3580 BP; 986 A; 823 C; 996 G; 775 T; 0 other;
alignment_scores:
Quality: 1583.00 Length: 1060
Ratio: 2.288 Gaps: 36
Percent Similarity: 65.283 Percent Identity: 36.415
alignment_block:
US-09-236-995D-2 x AA223799 ..
Align seg 1/1 to: AA223799 from: 1 to: 3580
6 LysAlaTrpLysAlaGluTyrAlaLysSerGlyArgAlaSerCysLysSe 22
158 AAGCTCTATCGAGTCGATAGCCCAAGAGCGGGCGGCTCTTGCAGAA 207
22 rCysArgSerProIleAlaLysAspGlnLeuArgLeuGlyLysMetVal 39
208 ATGACGAGGAGAGCATCCCAAGAGCTCGCTCCGATGCCATCATGTGTC 257
39 lAlaSerGlnPheAspGlyPheMetProMetTrpAsnHisAlaArgCys 55
258 AGTCCCATCTTTGATGGGAAAGTCCCACTGCTACCTCTCTCTCTGC 307
56 lIlePheSerLysLysAsnGlnIleLysSerValAsp...AspValGluGl 71

354 eLysSerGlnLysValLysLysProGluArgValLeuProMetSer. 370
 1174 CAAGAAATTTGAAGCTTAAAGACGAGACCGTATATTCGCCCAACAAA 1223
 371ProGluLysSerGlySerLysAla 378
 1224 GCGCCTCCGTGCGGCGCAGCCTCGCCCTCCACAGTCTCGGCTCGTGC 1273
 379 ThrGlnArgThrSerLeuLeuSerSerLysGlyLeuAspLysLeuArgPh 395
 1274 GCTGTGAACCTCTCTGCTTACGACAGATAAGCAATTATCCAAACATGA 1323
 395 eSerValValGly.....GlnSerLysGluAlaAlaAsnGluTrpI 409
 1324 CTGACTCTCGGGAAGCTGTCCCGGAACAGGATCAAGTGAAGGCCATGA 1373
 409 leGluLysLeu.....LysLeuAlaGlyAlaAsnPhETyrAlaArgVal 423
 1374 TTGAGAAACTCGGGGGAAGTTGACGGGGAGCGGCCAACAGGCTTCCCTG 1423
 424 ValLysAspIleAspCysLeuIleAlaCysGlyGluLeuAspAsnGluAs 440
 1424TGCATCAGCACCAAAAGAGGCTGGAAGAAGATGA 1458
 440 nAlaGluValArgLysAlaArgArgLeuLysIleProIleValArgLys 457
 1459 TAAGAAGATGAGGAAGTAAGGAAGCCACATCCGAGTTGTCTCTGAGG 1508
 457 LysTyrIleGlyGluCysValLysArgThrLys..... 467
 1509 ACTTCTCCAGCAGCTCTCGCCTCCACCAAGAGCCTTCAGGAGTTGTTC 1558
 468 CysCysHisLeuIleCysIleAsnTrpAsnAlaLeuGluSerSerLysG 484
 1559 TTAGCGCACATTTGTCTCC...CTTGGGGGCGAGGTGAAGCGACAGCC 1605
 484 y***ThrValThrValLysValLysGlyArgSerAlaCysSer***SerP 501
 1606 TGTTCAGATT...GTGGCCCCAGAGGGAAGTCAGGGGTGCGCTCTCCA 1652
 501 ro***ValCysLysAsnThrAlaHisIlePro***Trp..... 514
 1653 AA.....AAAGCAAGGGCGAGTCAAGGAGGAAGGTATCAACAAA 1693
 515 ...GluLysHisIleGlnCys***LeuLys..... 523
 1694 TCTGAAAGAGATGAATTAATCTTTAAGAGGAGGACGCTGTGTGATCC 1743
 524HisValLeu..... 526
 1744 TGATTCGTGAGTGGAAACACTCTCGCATGTCTGGAGAAAGGTGGGAAG 1793
 527ThrHis***Val.....CysThrGly 534
 1794 TCTTCAGTGCACCTTGGCCTGGTGGACATCCTTTAAGGAACCAACTCC 1843
 535 TyrTyrValLeuGlnIleIleGluGlnAspAspGlySerGluCysTyrVa 551
 1844 TACTACAAGCTCGACGCTCTGGAGGACGACAGGAAACACGCTATTGGAT 1893
 551 lPreArgLysTrpGlyArgValGlySerGluLysIleGlyGlyGlnLysL 568
 1894 ATTCAAGTCTCGGGGCGGTGTGGGTACG...GTCACTCGGTACCAACAA 1940
 568 euGluGluMet...SerLysThrGluAlaIleLysGluPheLysArgLeu 583
 1941 TGAACACAGATGCCGTCCAGGAGGATGCCATTCAGCACTTCATGAATTA 1990
 584 PheLeuGluLysThrGlyAsnSerTrpGluAlaTrpGluCysLysThrAs 600
 1991 TATCAAGAAAAACCGGACACGTTGGCACTCC.....AAAA 2028

PA (ICOS-) ICOS CORP.

XX Christenson E, Demaggio AJ, Goldman PS, McElligott DL;

XX WPI; 2001-025335/03.

DR P-PSDB; AAB47032.

XX New human poly(ADP-ribose) polymerase for treating inflammatory,
PT neurological, cardiovascular, or neoplastic tissue growth disorders,
PT such as, arthritis, encephalitis, myocardial ischemia, and leukocyte
PT metastasis

XX Example 3; Page 116-117; 129pp; English.

CC This sequence encodes the fusion protein PARP1A/PARP2B. This
CC protein contains amino acids 1-662 of hPARP1 fused upstream of
CC amino acids 230-583 of hPARP2. This sequence was amplified using
CC the primer sequences given in AAC85321-40 and AAC85342-51. The fusion
CC protein coding sequence was cloned in a baculovirus expression
CC vector for the expression of the fusion protein. The
CC protein of the invention, hPARP2, causes the covalent addition of
CC polymers of ADP-ribose to protein targets. hPARP2 activity is induced
CC in many instances of oxidative stress or during inflammation where
CC there is direct damage to the DNA. hPARP2 may be used to identify
CC antagonists which may be used to treat a human having a disorder
CC mediated by PARP2 activity, such as, inflammatory, neurological,
CC cardiovascular, or neoplastic tissue growth disorders. hPARP2 and
CC antibodies to it, can also be used to diagnose these conditions.

XX Sequence 3200 BP; 923 A; 745 C; 870 G; 662 T; 0 other;

alignment_scores:

Quality: 1518.50 Length: 1074
Ratio: 2.240 Gaps: 40
Percent Similarity: 63.128 Percent Identity: 35.754

alignment_block:

US-09-236-995d-2 x AAC85341

Align seg 1/1 to: AAC85341 from: 1 to: 3200

6 LysAlaTrpLysAlaGluTyrAlaLysSerGlyArgAlaSerCysLysSe 22
127 AACCTATFCAGTAGTCAGTACCCCAAGAGCGCGCCCTCTTGCAGAA 176
22 rCysArgSerProIleAlaLysAspGlnLeuArgLeuGlyLysMetValG 39
177 ATGTAGCAGAGACATCCCAAGGACTCGCTCGGATGGCCATCATGTGTC 226
39 InAlaSerGlnPheAspGlyPheMetProMetTrpAsnHisAlaArgCys 55
227 ATGCGCCATGTTGATGGAAAGTCCACACTGTGTACCACTTCTCTGTC 276
56 IlePheSerLysLysAsnGlnIleLysSerValAsp... AspValGluG1 71
277 TTCTGGAAGGTGGCCACTCCATCCGACCCCTGACGTTGAGGTGGATGG 326
71 yIleAspAlaLeuArgTrpAspAspGlnGluLysIleArgAsnTyrVal. 87
327 GTTCTCTGAGTCTCGGTGGATGACCAAGAGTCAAGAAAGTCAAGAACACCGG 376
88GlySerAlaSerAlaGlyThrSerSerThrAlaAlaProPro 101
377 AAGTGGAGGAGTACAGGCAAGAGGAGGATGGAAATTTGGTACCAAGGCA 426
102 GluLysCys.....ThrIleGluIleAlaProSerAlaArgth 114
427 GAGAAGACTCTGGGTGACTTTGGCAGCAGAGTATGTCAAGTCCCAACAGAAG 476
114 rSerCysArgArgCysSerGluLysIleThrLysGlySerValArgLeus 131
477 TACGTGAAGGGGTGTATGGAGAAGATAGAAAAGGGCCAGGTGCGCTGT 526

131 erAlaLysLeu...GluSerGluGlyProLys...GlyIle.....Pro 143
527 CCAAGAGATGGTGGACCCCGGAGGACACACGCTAGCATGATTGACCGC 576
144 TrpTyrHisAlaAsnCysPhePheGlu.....ValSe 154
577 TGTATACCATCCAGGCTCTTTGTCAAGAACAGGAGGAGCTGGTTTCGG 626
154 rProSerAlaThrValIleLysPheSerGlyTrpAspThrLeuSerAspG 171
627 GCCCGAGTACAGTCCGAGTCAGCTCAAGGCTTCAGCCCTCTTGTACAG 676
171 luAspLysArgThrMetLeuAspLeuValLysLysAsp.....Val 184
677 AGGATAAA.....GAAGCCCTGAAGAGCAGCTCCAGGAGTTC 714
185 GlyAsnAsnGluGlnAsnLysGlySerLys..... 194
715 AAGAGTGAAGGAAAGAGTAAAGCGATGAGTGGATGGATGGATCAAGT 764
195ArgLysLysSerGluAsnAspIleAspSerTyrLysSerAlaArgL 210
765 GGCAGAAAGAAATCTTAAAAAGAAAAAGAC.....AAGGATAGTAAGC 808
210 euAspGluSerThrSerGluGlyThrValArgAsnLysGlyGlnLeuVal 226
809 TTGAAAAAGCC..... 819
227 AspProArgGlySerAsnThrSerSerAlaAspIleGlnLeuLysLeu 243
820CTAAA 824
243 SGLuGlnSerAspThrLeuTrpLysLeuLysAspGlyLeuLysThrHisv 260
825 GGTCTCAGAACGACCTGATCTGGACATCAAGGAGCAGCTAAAGAAAGTGT 874
260 alSerAlaAlaGluLeuArgAspMetLeuGluAlaAsnGlyGlnAspThr 276
875 GTTCAACTAATGACCTGAAGGAGCTACTCATCTTCAACACGACCAAGTG 924
277 SerGlyProGluArgHisLeuLeuAspArgCysAlaAspGlyMetLeuph 293
925 CCTTCTGGGGAGTCGGCGATCTTGGACCGAGTAGCTGATGCTGCTGTT 974
293 eGlyAlaLeuGlyProCysProValCysAlaAsnGlyMetTyrTyrA 310
975 CGGTGCCCTTCTTCCCTCGAGGATGCTCGGTCAGCTGCTTCAAGA 1024
310 snGlyGlnTyrGlnCysSerGlyAsnValSerGluTrpSerLysCysThr 326
1025 GCGATGCTATTACTGCACTGGGAGCTGCTGCTGACCAAGTGTATG 1074
327 TyrSerAlaThrGluProValArgValLysLysLysTrpGlnIleProHi 343
1075 GTCAAGACACAGACACCAACCGG.....AAGGAGTGGTAAACCCCA.. 1116
343 sclyThrLysAsnAspTyrLeuMetLysTrpPheLysSerGlnLysVal 360
1117AAGAAATCCAGAAATCTCTTACCTCAAGAAATTTGAAGTTA 1159
360 ysLysProGluArgValLeuProMetSer..... 370
1160 AAAGCAGGACCGTATATTTCCCAAGAAACACGAGCGCTCCGTGGCGGC 1209
371ProGluLysSerGlySerLysAlaThrGlnArgThrSerle 384
1210 ACGCTCGCCCTCCACAGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1259
384 uLeuSerSerLysGlyLeuAspLysLeuArgPheSerValValGly.... 399
1260 TTCAGCAGATTAAGCCATTATCCACATGAAGATCTTCTGACTCTCGGGAAGC 1309

```

400 .....GlnSerLysGluAlaAlaAsnGluTrpIleGluLysLeu..... 412
413 TGTCCCGGAACAAAGGATGAAGTGAAGGCATGATTGAGAAATCGGGGG 1359
414 LysLeuAlaGlyAlaAsnPhetYrAlaArgValLysAspIleAspCy 429
415 AAGTTGACGGGACCGCCACAAAGCTTCCCTG.....TG 1394
429 sLeuIleAlaCysGlyGluLeuAspAsnGluAlaGluValArgLysA 446
435 CATCAGCACCAAAAGAGGTGGAAAGATGAATGAAGATGGAGGAAG 1444
446 laArgArgLeuLysIleProIleValArgGluGlyIleGlyGluCys 462
447 TAAAGGAAGCAACATCCAGTGTGTCTGAGGACTTCTCCAGGACGTC 1494
463 ValLysArgThrLys.....CysCysHisLeuIleCy 473
495 TCGCCCTCCACCAAGAGCCTTCAGGAGTGTCTTACGCCACATCTGTG 1544
473 sIleAsnTrpAsnAlaLeuGluSerSerLysGly***ThrValThrVal 490
491 C...CCTTGGGGGACAGAGTGAAGGACAGCTGTCTTCAAGTT...GTGG 1588
490 ysValLysGlyArgSerAlaCysSer***SerPro***ValCysLysAsn 506
492 CCCAAGAGGGAAGTCAGGGCTGCCTCTCCAAA.....AAAGC 1629
507 ThrAlaHisIlePro***Trp.....GluLysHisIleG 519
509 AAGGGCCAGGTCAAGGAGGAGGTATCAACAATCTGAAGAGAGATGAA 1679
519 nCys***LeuLys..... 523
1680 ATTAACCTCTTAAAGAGGAGGACAGCTGTGGATCCTGATCTGACAGTGAAC 1729
524 .....HisValLeu..... 1730
529 His***Val.....CysThrGlyTyrTrpValLeuGlnI 540
540 eIleGluAlaAspAspClySerGluCysTyrValPheArgLysTrpGly 557
1830 TCTGGAGGACCAAGAAACAGGTATGGATATTACAGTCTCTGGGGCC 1879
557 rgValGlySerGluLysIleGlyGlnLysLeuGluMet...Ser 572
1880 GTGGGTAGC...GTGATCGGTAGCAACAACTGGAACAGATGCCGTCC 1926
573 LysThrGluAlaIleLysGluPheLysArgLeuPheLeuGluLysThrG 589
1927 AAGGAGGATGCCAATGTAGACATTCATGAAATATATGAAGAAAAACCG 1976
589 yAsnSerTrpGluAlaTrpGluCysLysThrAsnPheArgLysGlnPro 606
1977 GAACGCTTGGCACTCC.....AAAATTTCCAGCAAGATCCCA 2014
606 lyArgPheTyrProLeuAspValAspTyrGly.....Val 617
2015 AAAAGTTCTACCCCTCGAGATTCAGTATGCCCAGGATGAAGAGGACGT 2064
618 LysLysAla.....ProLysArgLysAspIleSerGluMetLysSe 631
2065 AAGAAAGCTGACAGTAATCTTGGCACCACAAACGGCTTCCAGAGTCACAG 2114
631 rSerLeuAlaProGlnLeuLeuGluLeuMetLysMetLeuPheAsnVal 648
2115 AGATCTTCGGGTACAG.....GAGTTAATAAAGTTGATCTGTAATGTC 2158
648 luThrTyrArgAlaAlaMetMetGluPheGlu***AsnMetSerGluMet 664

```

```

2159 AGCCATGGGAAGAAATGATGTAATGAAGTATAATACCAAGAAAGCC 2208
665 ProLeuGlyLysLeuSer*****AsnIleGlu***GlyPheGluAla 681
2209 CCTTGGGAAGCTGACAGTGGCGCAATCAAGCGAGGTACAGCTCTCT 2258
681 uThr*****LeuPheGluGlyHisArg***SerSerThrGlyL 698
2259 TAAGAAGATTGAGATTGTATTCGG.....GCTGGCC 2290
698 eu***GluLysAla***LeuLeu*****PheSerLeuLeu 714
2291 AGCATGGACGAGCTCTCATGGAAGCATGCAATGAATTTACACAGGATT 2340
715 SerLeuLeuPheIleLeu.....IleLeuTyrGlyMetArgMe 727
2341 CCGCATGACTTGGACCTCCGCTACTCTCCACTAATCCGGACACAGAGA 2390
727 tIleSerTyrSerLysAlaLysMetLeuGluAlaLeuGlnAspIleGlu 744
2391 ACTGTCA...GAAAAAATAACAATTTACTAGAGGCTTTGGGAGACATTG 2437
744 leAlaSerLysIleValGlyPheAspSerAspSerGlu...SerLeu 759
2438 TTGCTATTAGCTGTGTAACACAGAGCTACAAGGCCAGACACACCCATT 2487
760 AspAspLysTyrMetLysLeuHisCysAspIleThrProLeuAlaHis 776
2488 GACCAACACTATAGAACCCTACATTTGCGCTTGGCCCTTGACCATGA 2537
776 pSerGluAspTyrLysLeuIleGluGlnTyrLeuLeuAsnThrHisAla 793
2538 AAGTTACGAGTTCAAAGTGATTTCCAGTACCTACAATCTACCATGCTC 2587
793 roThrHisLysAspTrpSerLeuGluLeuGluValPheSerLeuAsp 809
2588 CCACACAGCGACTATACCATGACTTCTGCTGGATTGTTGAAGTGGAG 2637
810 ArgAspGlyGluLeuAsnLysTyrSerArgTyrLysAsnAsnLeuHis 826
2638 AAGGATGCTGAG.....AAAGAAGCCTTCAGAGAGGACCTTCATAA 2678
826 nLysMetLeuLeuTrpHisGlySerArgLeuThrAsnPheValGlyIle 843
2679 CAGGATGCTTCTATGGCATGTTCCAGATGAGTAACCTGGGTGGGAATCT 2728
843 euSerGlnGlyLeuArgIleAlaProGluAlaProValThrGlyTyr 859
2729 TGAGCCATGGGCTTCGAATTCGCCACCTGAAGCTCCCATCACAGGTAC 2778
860 MetPheGlyLysGlyLeuTyrPheAlaAspLeuValSerLysSerAlaG 876
2779 ATGTTTGGAAAGAAATCTACTTTGCTGACATGTCTTCCAAGAGTGC 2828
876 nTyrCysTyrValAspArgAsnAsnProValGlyLeuMetLeuSerG 893
2829 TTACTGCTTGTGCTCTCGCCCTAAAGAAATACAGACTGCTGCTTATCAG 2878
893 luValAlaLeuGlyAspMetTyrGluLeuLysLysAlaThr...SerMet 908
2879 AGGTAGCTCTAGTCTAGTGTATTAATGAACACTAGAGGCCAATCTAAG 2928
909 AspLysProArgGlyLysHisSerThrLysGlyLeuGlyLysThrVa 925
2929 GAAGGATTCTCTCAAGGTAAACATACAGCAAGGGCTGGGCAAGATGCC 2978
925 lProLeuGluSerGluPheValIleTrpArgAspValValValProC 942
2979 TCCAGTTCTGCGCCACTCTGTCACCTCGAATGGAGT...ACAGTCCCAT 3025
942 ysGlyLysProValProSerSerIleArgSerSerGlu.....Leu 955

```

3026 TAGGACGACGAAGTGACACAGGAATTTCTGAATCCAGATGTTATACCTC 3075
 956 MetTyrAsnGluTyrIleValTyrAspThrSerClnValLysMetGlnPh 972
 3076 AACTACAATGAATATATTGTATATAACCCCAACGAGTCCGTATGCGGTA 3125
 972 eLeuLeuLysValArgPheHis 979
 3126 CCTTTAAAGGTTTCAGTTTAAT 3147

seq_name: /SDS2/gcgdata/geneseq/geneseq/NA1995.DAT: AAT04221

seq_documentation_block:

ID AAT04221 standard; DNA; 5345 BP.

AC AAT04221;

DT 31-MAR-1996 (first entry)

DE 5.3 kb EcoRI fragment containing poly (ADP-ribose)-polymerase gene.

XX DNA primer; PCR; polymerase chain reaction;

KW poly (ADP-ribose) polymerase; cancer diagnosis; chromosome-13;

KW pseudogene; Burkitt's lymphoma; B-follicular cell lymphoma;

KW small cell lung carcinoma; colorectal carcinoma; tumor;

KW breast carcinoma.

XX Homo sapiens.

OS US449605-A.

PN 12-SEP-1995.

XX 14-OCT-1988; 88US-0257696.

XX 06-APR-1993; 93US-0044618.

PR 14-OCT-1988; 88US-0257696.

XX (GEOU) UNIV GEORGETOWN.

XX Cherney B, Lyn D, Smulson ME;

PI WPI; 1995-327692/42.

XX Allele-specific primers for detecting pre-disposition to cancer -

DR caused by a deletion in chromosome 13 pseudo-gene for poly

PT (ADP-ribose) polymerase

XX Disclosure; Page 41-46; 47pp; English.

PS This sequence corresponds to a 5.3 kb EcoRI fragment that contains

XX the poly (ADP-ribose)-polymerase pseudogene.

CC Sequence 5345 BP; 1598 A; 1142 C; 1396 G; 1209 T; 0 other;

XX

alignment_scores:

Quality: 1237.00 Length: 1128

Ratio: 1.811 Gaps: 49

Percent Similarity: 60.550 Percent Identity: 33.511

alignment_block:

US-09-236-995D-2 x AAT04221 ..

Align seg 1/1 to: AAT04221 from: 1 to: 5345

6 LysAlaTrpLysAlaGluTyrAlaLysSerGlyArgAlaSerCysLysSe 22

791 AAGTCTACTGAGTGTACACCAAGAGCGGCGCACCTCTTGCAGAA 840

22 rCysArgSerProIleAlaLysAspGlnLeuArgLeuGlyLysMetValG 39

841 ATGCAGCGAGAGCATCCCAAGGAGTCCGTCGATGCGCCATCATGTGTC 890

39 InAlaSerGlnPheAspGlyPheMetProMetTyrAsnHisAlaArgCys 55
 891 AATAGTCCATGTT.GATGGAGAAGTCCCACACTGGTACCACCTTCTCTCTGC 939
 56 IlePheSerLysLysAsnGlnIleLysSerValAsp...AspValGluG1 71
 940 TTGCAGAGAGTGGCGCACTCCCATGGCGCACCTGAGTGGATGG 989
 71 yIleAspAlaLeuArgTyrAspGlnGluLysIle.....ArgA 85
 990 ATTCTCTGAGCTTTGGTGGATGACACGAGAGAAAGTCAAGAGACACGTG 1039
 85 snTyrValGlySerAlaSerAlaGlyThrSerSerThrAlaAlaProPro 101
 1040 ACGCTGGAGGAGTGACGGGCAAAAGGCCACCAATGGAATTGGTAGCAAGCA 1089
 102 GluLysCys.....ThrIleGluIleAlaProSerAlaArgTh 114
 1090 GAGAAGACGCTGGTGACTTTGCAGCAGAGTAGGCCAAGTCCAGCAGAG 1139
 114 rSerCysArgArgCys.SerGluLysIleThrLysGlySerValArgLeu 130
 1140 CAAGTCACAGAGTGTATGGAGAGATAGAAAAGGCCAGATGCGCTA 1189
 131 SerAlaLys...LeuGluSerGluGlyProLys...GlyIle.....Pr 143
 1190 TCCAAGAAGATGCTGGACCTGAGAAGCCTCAGTAGGCATGATTGACCC 1239
 143 oTrpTyrHisAlaAsnCysPhePheGlu.....Vals 154
 1240 CTGGTACCCACCGGCTGCTTGTCAAGAACAGGAGGAGCTGGGTTC 1289
 154 erProSerAlaThrValGluLysPheSerGlyTyrAspThrLeuSerAsp 170
 1290 AGCTGAGTACAGTGGAGTCAAGGCTTCAGCCTCTCTCGCTGCA 1339
 171 GluAspLysArgThrMetLeuAspLeuValLysLysAspValGlyAsnAs 187
 1340 GAGGATAAAGAAACCTG.....AAGAAGCAGCTCCAGGAGT 1377
 187 nGluGlnAsnLysGlySerLysArgLysLysSerGluAsnAspIleAsp 204
 1378 CAAG.....AGTGAGAAAGAGAGAACGCGGATGAGGTGGATG 1415
 204 erTyrLysSerAlaArgLeuAspGluSerThrSerGluGlyThrValArg 220
 1416 GA.....GTGGATGA.....GTGGCCAAG 1435
 221 AsnLysGlyGlnLeuValAspProArgGly.SerAsnThrSerSerAlaA 237
 1436 AAGAAATCTAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1484
 237 spIleGlnLeuLysLeuLysGluLysSerAspThrLeuTrpLysLeuLys 253
 1485CTAAGGCCACGAGACGCTGATCTGGAAACATCAAG 1520
 254 AspGlyLeuLysThrHisValSerAlaAlaGluLeuArgAspMetLeuG1 270
 1521 GACGAGCTAATGAAGTGTCTTATTAATGACCTGAGGAGATGCTCAT 1570
 270 uAlaAsnGlyGlnAspThrSerGlyProGluArgHisLeuLeuAspArgC 287
 1571 CTTCAACAGGACGAGTGCCTTCGGGAGTGGCGGATCTTGGACCGAG 1620
 287 ysAlaAspGlyMetLeuPheGlyAlaLeuGlyProCysProValCysAla 303
 1621 TAGCTGACAGCATGCTGGTGGCTCTCTCCCTGTGAGGAATGCTCA 1670
 304 AsnGlyMetTyrTyrTyrAsnGlyGlnTyrGlnCysSerGlyAsnValSe 320
 1671 GGTCAAGTGTCTTCAAGACGACACTTATTACTGCACCGGAGACGTAC 1720

480 TCCTTACTCTTGGGAGCTCTCCCGAACAAGGATGAAGTGAAGCCACG 529
409 IleGluLysLeuLysLeuAlaGlyAlaAsnPheTyrAlaArgValValY 425
|||||
530 ATTGAGAACTC.....GGGGAAAGTTGTGGGACGGCCAAACA 570
425 sAspIleAspCysLeuIleAlaCysGlyGluLeuAspAsnGluAsnAlaG 442
|||||
571 GGCTCTCGTCATCAGCACTTAAAGAGAGGGTGGAAAAGAAAGAAAGA 620
442 luValArgLysAlaArgArgLeuLysIleProIleValArgGluGlyTyr 458
|||||
621 AGATGGGAAGTGAAGGAAGCAACACTGAGCTGTCTGAGGACTTC 670
459IleGlyGlu.CysValLysA 465
|||
671 CTCAGGACTTCTGCTCCACCAAGGGCTCCAGGAGTTGTCTCAGC 720
465 rgThr..... 466
|||||
721 GCACATATTGACCCCTGGGGGCGAGAGGTGAAGGCAGAGCCTGTGAAG 770
467LysCysCysHis..... 470
|||||
771 TCCTAGCCCCAAGAGGAAGTCAAGAGCTGTCTCTCCAAAAAAGCAAG 820
470 470
821 GGCCAGTCAAGGAGGAGGTATCAACAAATCTGAAAAGAGAATGAAATT 870
471LeuIleCysIleAsnTrpAsnAla 478
|||||
871 AACTCTTAAAGGAGGAGCAGCTGTGGATCCTGACTGTCTGCGAACAAC 920
479 LeuGluSerSer..... 482
|||||
921 CTGCGCATCTCTGGAGAAAGGTGAAGGCAGAGCCTGTTGAAGTCTGAGC 970
483LysGly***ThrValThrValLysValLysGlyArg. 494
|||||
971 CCCAAGAGGGAAGTCAAGAGCTGTGCTCTCCAAAAAAGCAAGGCCAAG 1020
494 494
1021 TCAAGGAGGAGGTATCAACAACATCAAAAGAGATGAATTAACCTCT 1070
495SerAlaCysSer***SerPro***ValCysLysAsnThrAlaHi 509
|||||
1071 AAAGGAGGAGCAGCTCTGGATCTGACTCTGCTGCTGGAACACTCTGCACA 1120
509 sIlePro***TrpGluLysHisIleGlnCys***LeuLysHisValL 526
|||||
1121 TGTTCTGGAGAAAGTGGGAAGGTCTCAGTGCCACCTCAGCCCTGGTG. 1169
526 eutThr***His***ValCysThrGly.....TyrTyrValLeuGln 539
|||||
1170GACGTCTGTAAGGAACCACTCTCTATTACAAGTGAAGT 1208
540 IleIleGluGlnAspAspGlySerGluCysTyrValPheArgLysTrpGl 556
|||||
1209 TTGCTGAAGGATGACAAGAAAGCAGGCATTTGGATATTCAAGTCTCGGA 1258
556 YArgValGlySerGluLysIleGlyGlnLysLeuGluMet...s 572
|||||
1259 CCCTGTGGGCACG...GTGATCGGTAGCAACAACCTGGACAGAGTCTCT 1305
572 erLysThrGluAlaIleLysGluPheLysArgLeuPheLeuGluLysThr 588
|||||
1306 CCAAGGAGGACACCATTTGAACACTTTCATGAAATATATGAAGAAAACTA 1355
589 GlyAsnSerTrpGluAlaTrpGluCysLysThrAsnPheArgLysGlnPr 605
|||||
1356 GGAAT.....GCTTGGCACTCCAAAA...TTCAAAAGTATCC 1391

605 oGlyArgPheTyrProLeuAspValAspTyrGlyVal...LysLysAlap 621
| : : : : :
1392 CAAAGAGTTTACCCCTGGAGATTGACTACGGCCAGGACAAAGAGCGG 1441
621 roLysArgLysAspIleSerGluMetLysSerSerLeuAlaProGln... 636
| : : : : :
1442 TGAAGAGCCGACAGTAATCTCTGGCACCAGTCCATGCTCCCAAGCCA 1491
637 LeuLeuGluLeuMetLysMetLeuPheAsnValGluThrTyrArgAlaAl 653
| : : : : :
1492 GTTCAGGACCTATCAAGATGATCTTTGATGCTGAAAGTATGCCAAGGC 1541
653 aMetMetGluPheGlu***AsnMetSerGluMetProLeuGlyLysLeu 670
| : : : : :
1542 CCATGGGGGTGAGATCAACCTT...CAGATGCCCTTGGGAAAGCTGA 1588
670 er****AsnIleGlu***GlyPheGluAlaLeuThr***** 686
| : : : : :
1589 GCAAAAGGCAATCCAGGCCGCTACTCCATCTCTCAGGTC..... 1628
687 LeuPheGluGlyHisArg***SerSerThrGlyLeu***GluLysAla** 703
| : : : : :
1629CAGCAGGTGTGTCTCCAGGGCAGCAGCAGCTCTCAGAT 1666
703 *LeuLeu*****PheSerLeuLeuSerLeuPheIleL 720
| : : : : :
1667 CCTG.....GATCTCTCAATCGCTTTTACATCC 1695
720 euIleLeu.....TyrGlyMetArg.....MetIleSer..... 729
| : : : : :
1696 TGATCCCCACGAGCTTTGGGATGAAGATCCTCTGCTCTCTGAAACAATGCA 1745
730TyrSerLysAlaLysMetLeuGluAlaLeuGlnAspIleGl 743
| : : : : :
1746 GACAGTGTGAGGCCAGGTAGAAATGCTGCACAACTCTGCTGACATGTA 1795
743 uIleAlaSerLysIleVal.....GlyPheAspSerAspSerAspGlu 758
| : : : : :
1796 GGTAGCTACGGTCTGCTCAGGGGAGGTCTCTCAGATAGCAGGAAGCACT 1845
758 erLeuAspAspLysTyrMetLysLeuHisCysAspIleThrProLeuAla 774
| : : : : :
1846 CCATCGATGTCAACTATGAGAAGCTCAAACTGACATTAAGTGGTGGTAC 1895
775 HisAspSerGluAspTyrLysLeuIleGluGlnTyrLeuLeuAsnThrHi 791
| : : : : :
1896 AGATTTCTGAAGAGCTGAGATCATCAGGAGTATGTTAAGACACTCA 1945
791 sAla.Pro...ThrHisLysAspTrpSerLeuGluLeuGluValPhe 806
| : : : : :
1946 TGCAACCAACCAACACACGATCATATGACTTGAAGTCAATTGATAGCTTT 1995
807 SerLeuAspArgAspGlyGluLeuAsnLysTyrSerArgTyrLysAsnAs 823
| : : : : :
1996 AAGTAGAGTGTGAAGAGAGTGCACACTACAAAGCCCTTTAAG...CA 2042
823 nLeuHisAsnLysMetLeuLeuTrpHisGlySerArgLeuThrAsnPheV 840
| : : : : :
2043 GCTTCATPACTGAAGTGTCTGTCATGCTGCTGAGGAGTGTCCAGGACCACTTG 2092
840 alGlyIleLeuSerGlnGlyLeuArgIleAlaProProGluAlaProVal 856
| : : : : :
2093 CTGGGATCTCTCCTGGTCTTTGGATAGCCCTGCTGCTGAAGCACTGTG 2142
857 ThrGlyTyrMetPheGlyLysGlyLeuPheAlaAspLeuValSerLy 873
| : : : : :
2143 ATGGCTACATGTTGGTAAAGTAGTCTATTTCGCTGATCTGTCTCAA 2192
873 sSerAlaGlnTyrCystyrValAspArgAsnAsnProValGlyLeuMetL 890
| : : : : :
2193 GAGTCCCAACGACTGCCATACATCTTAGAAGACCAATAGGGTTAATCC 2242

524HisValLeu..... 526
 1713 TGATTCTGGAGTGAACACACTCTCGCATGCTCTGGAGAAAGTGGGAAGG 1762
 527Thr***His***Val.....CysThrGly 534
 1763 TCITTCAGTGGCCACCTTGGCTGGTGCACATCGTTAAAGGAACCACTCC 1812
 535 TyrTyrValLeuGlnIleIleGluGlnAspGlySerGluCysTyrVa 551
 1813 TACTACAGCTGTCAGCTTCTGGAGGACGACAAAGAAACAGGATTGGAT 1862
 551 lPheArgLysTrpGlyArgValGlySerGluLysIleGlyGlnLysL 568
 1863 ATTCAAGTCTGGGGCGCTGGGTAGC...GTGATCGGTAGCAACAAC 1909
 568 euGluGluMet...SerLysThrGluAlaIleLysGluPheLysArgLeu 583
 1910 TGAACAGATGCCGTCCAAAGGAGGATGCCATGTGACACTTTCATGAAATTA 1959
 584 PheLeuGluLysThrGlyAsnSerTrpGluAlaTrpGluCysLysThrAs 600
 1960 TATGAAGAAAAACCGGGAACGCTTGGCACATCC.....AAAA 1997
 600 nPheArgLysGlnProGlyArgPheTrpTyrProLeuAspValAspTyrGly. 616
 1998 TTTCACGAAGTATCCCAAAAGTTCTACCCCTCGGAGATTGACTATGGCC 2047
 617ValLysLysAla.....ProLysArgLysAsp 625
 2048 AGGTGAAGAGGAGTGAAGAAGCTGCAGTAAATCTTGGCACCACAACGC 2097
 626 lle..... 626
 2098 GTTGAGACAAACACCTTTAGATTAGTTTTCAGCAGATGTCAGCGC 2147
 627SerGluMetLysSerSerLeuAlaP 635
 2148 TCCTCTGACAGCAGCATGCCCATCTGCTCTGCCCTCTTGTTCACAGC 2197
 635 roGlnLeuLeuGluLeuMetLysMetLeuPheAsnValGluThrTyrArg 651
 2198 CTCACAGTCTCAATGGTGTGAGA.....AGCCAGGAGCCACTGCA 2238
 652 AlaAlaMetMetGluPheGlu***AsnMetSerGluMetProLeuGlyLy 668
 2239 GATGCTCTCTTCAGGTCATCTAGCCCATCAAGCCTTCTGACAGCCAG 2288
 668 sLeuSer*****AsnIleGlu**GlyPheGluAlaLeuThr***** 685
 2289 C...AGTCTTGACAACTTATCTGGGAGTTCAGAACTGCTTCCAGTA. 2334
 685 *****LeuPheGluGlyHisArg***SerSerThrGlyLeu***GluLys 701
 2335GTTAGTTCAAGTGA...ACAGAGGT 2358
 702 Ala***LeuLeu*****pHeSerLeuLe 714
 2359 GCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAAC 2408
 714 uSerLeuLeuPheIleLeuIleuLeuTyrGlyMetArgMetIleSerTyrS 731
 2409 TCAATTCTGAAGAAATCTTGGAGCTTGAGCACCTAATGGATATATTTGACA 2458
 731 eLysAlaLysMetLeuGluAlaLeuGlnAspIle..... 742
 2459 GAGACAGATCACTTTGGATGTATTAGTTGAGATGGGCGACAAAGAGCTG 2508
 743 ...GluIleAlaSerLysIleValGlyPheAspSerAspSerAspGluSe 758
 2509 AAGGAGATTGGAATCAATCTTATGGACATAGGCACAAACTAATTAAGG 2558
 758 rLeuAsp.....AspLysTyrMetLysL 766

254 pGlyLeuLysThrHisValSerAlaAlaGluLeuArgAspMetLeuGluA 271
 858 CGAGCTAAAGAAAGTGTGTTCAACTAATACCTGAAGGAGCTACTCATCT 907
 271 laAsnGlyGlnAspThrSerGlyProGluArgHisLeuLeuAspArgCys 287
 908 TCAACAAGCAGCAAGTGTCTTCTGGGAGTCCGGCATCTTGGACCGAGTA 957
 288 AlaAspGlyMetLeuPheGlyAlaLeuGlyProCysProValCysAlaAs 304
 958 GCTGATGGCATGTGTGCGTCCCTCTTCCCTCGGAGGATGTCGCGG 1007
 304 nGlyMetTyrTyrAsnGlyGlnTyrGlnCysSerGlyAsnValSerG 321
 1008 TCAGCTGGTCTTCAAGAGCGATGCTTACTGCACTGGGGAGCTCACTG 1057
 321 luTrpSerLysCysThrTyrSerAlaThrGluProValArgValLysLys 337
 1058 CCTGGACCAAGTGTATGGTCAAGACACAGACACCCCAACCG.....AAG 1101
 338 LysTrpGlnIleProHisGlyThrLysAsnAspTyrLeuMetLysTrpPh 354
 1102 GATGGGTACCCCA.....AGGAATCCGAGAAATCTCTTACCT 1142
 354 eLysSerGlnLysValLysLysProGluArgValLeuProPrometSer. 370
 1143 CAAGAAATTCAGGTAAAGAACGAGGACCGTATATTTCCCCCGAGAACCA 1192
 371ProGluLysSerGlySerLysAla 378
 1193 GCGCCTCCGTGGCGGCCACCCCTCCGCCCTCCACAGCCTCGGCTCTGCT 1242
 379 ThrGlnArgThrSerLeuLeuSerSerLysGlyLeuAspLysLeuArgPh 395
 1243 GCTGTGAATCTCTCTGCTTCCAGCAGATAGCCATTATCCAAACATGAAGAT 1292
 395 eSerValValGly.....GlnSerLysGluAlaAlaAsnGluTrpI 409
 1293 CTTGACTCTCGGAAGCTGTCCCGGACCAAGGATGAAGTGAAGGCCATGA 1342
 409 leGluLysLeu.....LysLeuAlaGlyAlaAsnPheTyrAlaArgVal 423
 1343 TTGAGAACTCGGGGGAAGTTGACGGGAGCGCCCAACAGGCTTCCCTG 1392
 424 ValLysAspIleAspCysLeuIleAlaCysGlyGluLeuAspAsnGluAs 440
 1393TGCATCAGCACCAAAAAAGGAGGTGGAAGATGAA 1427
 440 nAlaGluValArgLysAlaArgLeuLysIleProIleValArgGluG 457
 1428 TAAGAAATGGAGGAAGTAAAGGAAGCAACATCCGAGTTGTGTGAGG 1477
 457 lyTyrIleGlyGluCysValLysArgThrLys..... 467
 1478 ACTTCTCCAGGAGCTTCCGCCCTCCACCAAGAGCCTTCAGGAGTTGTTTC 1527
 468 CysCysHisLeuIleCysIleAsnTrpAsnAlaLeuSerSerLysG1 484
 1528 TTAGCGCACATCTTGTCC...CCTTGGGGGACAGAGTGAAGCGAGAGCC 1574
 484 y***ThrValThrValLysValLysGlyArgSerAlaCysSer***SerP 501
 1575 TGTGGAAGTT...GTGGCCCCCAAGGAGGAGTCAAGGGCTGCGCTC1CCA 1621
 501 ro***ValCysLysAsnThrAlaHisIlePro*****Trp..... 514
 1622 AA.....AAAAGCAAGGCCAGCTCAAGGAGGAAGGTATCAACAAA 1662
 515 ...GluLysHisIleGlnCys***LeuLys..... 523
 1663 TCTGAAAGAGAAATTAATCTTAAAGGAGGAGGACGCTGTGGATCC 1712

```

2559 AGTCGAGAGACTTATCTCCGGACAAACAGGCTTTAAACCAATATTTAACTT 2560
      :::::      ::::  |||::: |
766 euHisCysAsp.....lleThrProLeuAlaHisAspSer 777
      ||:::  ::::  |||:::  ::::
2609 TGAACACCTCTGGTAGTGGAACAATCTTATAGATCTGTCTCTCTGATGAT 2658
      :::::  ::::  |||:::  ::::
778 GluAspTyrLysLeuIleGluGlnTyrLeuLeuAsnThrHisAlaProTh 794
      :::::  ::::  |||:::  ::::
2659 AAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAGTACA...GTTCCAGA 2705
      :::::  ::::  |||:::  ::::
794 rhisLysAspTrpSer.....GTTCCAGA.....L 800
      |||:::  |||:::  |||:::  |||:::
2706 GCACAGATGGAGTCTCATCGAGTGGAACTTCAACAGATACAAATATTC 2755
      :::::  ::::  |||:::  ::::
800 euGluLeuGluValPheSerLeuAspArgAspGlyGluLeuAsnLys 816
      |||:::  ::::  |||:::  ::::
2756 TCAAGATTCAAGAGGTT.....TGTAACAAG 2781
      :::::  ::::  |||:::  ::::
817 .....TyrSerArgTyr.....LysAs 822
      :::::  ::::  |||:::  ::::
2782 AAATATGGGAAGATACATCACTACCGGAGAAAGAAGATTCTTCAAGAAAA 2831
      :::::  ::::  |||:::  ::::
822 nasNLeuHis...AsnLysMetLeuLeuTrpHisGlySerArgLeuThrA 838
      ||:::  |||:::  ::::  |||:::  ::::
2832 CCACAACCATGCCAATGAACGAATCTATTTCATGGGTCTCTCTTTGTGA 2881
      :::::  ::::  |||:::  ::::
838 snPheValGlyLeuLeuSerGlnGlyLeuArgIleAlaProGluAla 854
      ||  :::::  ::::  |||:::  ::::
2882 AT.....GCAATTATCCACAAAGGCTTT.....GATGAAGAGGCAT 2916
      |||  :::::  ::::  |||:::  ::::
855 ProValThrGlyTyrMetPheGlyLysGlyLeuTyrPheAlaAspLeuVa 871
      |||  :::::  ::::  |||:::  ::::
2917 CGGTACATAGGTGCTATGTTTGGAGCTGGCATTTATTTCTCGAAACACT 2966
      |||  :::::  ::::  |||:::  ::::
871 IserLysSerAlaGlnTyr..... 877
      :::::  ::::  |||:::  ::::
2967 TTCCAAAAGCAATCAATATGTATGGAATGGAGGAGGTACTGGGTGTC 3016
      :::::  ::::  |||:::  ::::
878 .....CystyrValAspArgAsnAsnProValGly 887
      |||:::  |||:::  |||:::  |||:::
3017 CAGTTCAACAAGACAGATCTTGTTACATTTGCCACAGGCAG..... 3057
      :::::  ::::  |||:::  ::::
888 LeuMetLeuLeuSerGluValAlaLeuGlyAspMetTyrGluLeuLysLy 904
      |||:::  ::::  |||:::  ::::
3058 ..CTGCTCTTTTGC CGGGTAAACCTTGGGAAAGTCTTCTCGCAGTTCCAG 3104
      :::::  ::::  |||:::  ::::
904 salAthrSerMetAspLysProProArgGlyLysHisSerThrLysGlyL 921
      |||:::  ::::  |||:::  ::::
3105 TCCAATGAAATGGCACAATCTCTCCAGGTCACTCACTAGTCACCTGGGT. 3153
      :::::  ::::  |||:::  ::::
921 euGlyLysThrValProLeuGluSerGluPheValLysTyrArgAspAsp 937
      :::::  ::::  |||:::  ::::
3154 .....AGGCCAGT 3162
      :::::  ::::  |||:::  ::::
938 ValValValProCysGlyLysProValProSerSerIleArgSerSerGl 954
      |||  :::::  ::::  |||:::  ::::
3163 GTA.....AATGG 3170
      :::::  ::::  |||:::  ::::
954 uLeuMetTyrAsnGluTyrIleValTyrAsnThrSerGlnValLysMetG 971
      |||  :::::  ::::  |||:::  ::::
3171 CCTAGCATTTACGTAAATATGTTATTACAGAGGAGACAGGCTTATCCCG 3220
      :::::  ::::  |||:::  ::::
971 InPheLeuLeu 974
      :::::  |||:::
3221 AGTATTTAAAT 3231
      :::::  |||:::

```

Align seg 1/1 to: AAF59996 from: 1 to: 1566

```
519 GlnCys***LeuLysHisValLeuThr***His***ValCysThrGly... 534
:::||||| ||| ||| ||||| ||| |||
103 GAGTGTACAGCAAG...GTGGGAGAGCTCATGTGTTGTAAGGAA 149
534 .....
150 TGATGCTATGATGTCATGCTAAATCAGACCAATCTCCAGTTCAACAACA 199
535 .....TyrTyrrValLeuGlnIleLeuGlnAspAspGlySerGluCys 549
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
200 ACAAGTACTACTGATTCAGCTATAGAGATGATGCCAGAGAACTTC 249
550 TyrValPheArgLysTrpGlyArgValGlySerGluLysIleGlyGly 566
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
250 AGTGTGTTGGATGAGATGGGGCGAGTTGG.....AAATGGGACAGCA 293
566 nLysLeuGluGluMetSer.....LysThrGluAlaIleLysGluPheL 581
:::||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
294 CAGCCTGTGTGCTTTCAGGCAATCTCAACAAGGCCCAAGGAATCTTTC 343
581 ysArgLeuPheLeuGluLysThrGlyAsnSerTrpGluAlaTrpGluCys 597
:::||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
344 AGAAGAAATCTCTTGCACAAAACGAAAAACAATGGGAAGATCGAGAA... 390
598 LysThrAsnPheArgLysGlnProGlyArgPheTyrProLeuAspValAs 614
:::||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
391 .....AGTTTGGAGAGGTGCTTGAAGAAATATGATATGTACAGATCGA 434
614 pTyrGlyValLysLysAlaProLysArgLysAspIleSerGlu.....M 629
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
435 CTATCCCACTACTCAGGATGAGAGGAACAAAGAAAGAGGAATCTC 484
629 eLysSerSerLeuAlaProGln.....LeuLeuGlu 639
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
485 TTAATATCTCCTTGAAGCCAGAGTACAGCTAGATCTTCGGGTACAGGAG 534
640 LeuMetLysMetLeuPheAsnValGluThrTyrArgAlaAlaMetMetG 656
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
535 TTAATAAAGTTGATCTGTAATGTTGAGGCCATGGAAGAAATGATGAGA 584
656 pHeGlu***AsnMetSerGluMetProLeuGlyLysLeuSer*****A 673
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
585 AATGAAGATGATTAACCAAGAAAGCCCACTGGGAAGCTGACAGTGGCAC 634
673 snIleGlu***GlyPheGluAlaLeuThr*****LeuPheGlu 689
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
635 AAATCAAGGCAGGTTACCAGTCTCTTAAGAAGATTGAGGATTGATTTCG 684
690 GlyHisArg***SerSerThrGlyLeu***GluLysAla***LeuLeu** 706
:::||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
695 .....CTGGCCAGCATGCGACAGCTCTCATGGAAGC 716
706 *****PheSerLeuLeuSerLeuLeuPheIleLeu..... 720
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
717 ATGCAATGAATTCATACAGGATTCGCGATGACTTTGGACTCCGCTACTC 766
721 .....IleLeuTyrGlyMetArgMetIleSerTyrSerLysAlaLysMet 735
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
767 CTCCTACTACCGGACACAGGAAGCACTGTCA...GAAAAAATACAAATTA 813
736 LeuGluAlaLeuGlnAspIleAlaSerLysIleValGlyPheAs 752
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
814 CTAGAGGCTTTGGGAGACATTTGAATGCTATTAGCTGGTGAACACAGA 863
752 pSerAspSerAspGlu...SerLeuAspAspLysTyrMetLysLeuHisC 768
:::||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
864 GCTACAAAGCCCAAGAACCCCACTTGGACCAACACTATAGAACCTACATT 913
768 ysAspIleThrProLeuAlaHisAspSerGluAspTyrLysLeuIleGlu 784
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
914 GTGCCTTGGCCGCCCTTGACCATGAAGTTATGAGTTCAAAAGTATTTC 963
```

```
785 GlnTyrLeuLeuAsnThrHisAlaProThrHisLysAspTrpSerLeuG 801
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
964 CAGTACCTACAAATCTACCATGCTCCACACACAGCGACTATACCATGAC 1013
801 uLeuGluGluValPheSerLeuAspArgAspGlyGluLeuAsnLysTyrS 818
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1014 CTTGCTGGGATTTGTTGAAGTGAGAGGATGGTGAG.....AAAG 1054
818 erArgTyrLysAsnAsnLeuHisAsnMetLeuLeuTrpHisGlySer 834
:::||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1055 AACCTTCAGAGAGGACCTTCATACAGAGATGCTCTATGCGCATGGTTC 1104
835 ArgLeuThrAsnPheValGlyIleLeuSerGlnGlyLeuArgIleAlaPr 851
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1105 AGGATGAGTAACCTGGCTGGGAATCTTGAAGCATGGCTTCGAATTCGCC 1154
851 oProGluAlaProValThrGlyTyrMetPheGlyLysGlyLeuTyrPheA 868
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1155 ACCTGAAGCTCCCATCACAGGTTCATGTTGGAAAGGAATCTACTTTTG 1204
868 laAspLeuValSerLysSerAlaGlnTyrCystyrValAspArgAsnAsn 884
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1205 CTGACATGCTTCCAAGAGTGCCAATCTGCTTTCCTCTCCCTTAAG 1254
885 ProValGlyLeuMetLeuSerGluValAlaLeuGlyAspMetTyrG 901
:::||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1255 AATACAGGACTGCTGCTCTATCAGAGGTAGCTTAGGTGAGTGAATGA 1304
901 uLeuLysLysAlaThr...SerMetAspLysProProArgGlyLysHisS 917
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1305 ACTACTAGAGGCCAATCTTAAGGCCGAAGGATTGCTTCAAGGTAAACATA 1354
917 erThrLysGlyLeuGlyLysThrValProLeuGluSerGluPheValLys 933
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1355 GCACCAAGGGCTGGCAAGATGGCTCCAGATGCTGCCCATCTCGGCACC 1404
934 TrpArgAspAspValValValProCysGlyLysProValProSerSerI 950
:::||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1405 CTCAATGGGAGT...ACAGTGCCATTAGGACCAAGTACACAGGAAT 1451
950 eArgSerSerGlu.....LeuMetTyrAsnGluTyrIleValTyrA 964
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1452 TCTGAATCCAGATGGTTATACCTCAACTACAATGAATATATGTATATA 1501
964 snThrSerGlnValLysMetGlnPheLeuLysValArgPheHis 979
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1502 ACCCAACAGGTCCTGATCGGTACCTTTTAAAGGTTTCAGTTTAAT 1548
seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT: AAC85303
```

seq_documentation_block:

ID AAC85303 standard; cDNA; 1814 BP.

XX AAC85303;

XX 29-MAR-2001 (first entry)

XX hpap2 cDNA.

DE

XX

KW Human; poly(ADP-ribose) polymerase; hPARP2; oxidative stress; ARDS;

KW inflammation; ischaemic stroke; hemorrhagic shock; myocardial ischemia;

KW infarction; cerebral vasospasm; rheumatoid arthritis; osteoarthritis;

KW gouty arthritis; spondylitis; Behcet's disease; sepsis; septic shock;

KW endotoxic shock; gram negative sepsis; gram positive sepsis; trauma;

KW toxic shock syndrome; multiple organ injury syndrome; vasculitis;

KW hemorrhage; conjunctivitis; uveitis; thyroid-associated ophthalmopathy;

KW eosinophilic granuloma; asthma; chronic bronchitis; allergic rhinitis;

KW chronic obstructive pulmonary disease; silicosis; reperfusion injury;

KW pulmonary sarcoidosis; pleurisy; alveolitis; pneumonia; myocardium;

KW bronchiectasis; pulmonary oxygen toxicity; keloid formation; brain;

KW scar tissue formation; atherosclerosis; systemic lupus erythematosus;

KW autoimmune thyroiditis; multiple sclerosis; Reynaud's syndrome;

736 LeuGluAlaLeuGlnAspIleuLeuAlaSerLysLysIleValGlyPheAs 752
|||||
1062 CTAGAGGCTTTGGGAGACATTGCTATTAGCTGGTGAACACAGA 1111
752 pSerAspSerAspGlu...SerLeuAspAspLysTyrMetLysLeuHisC 768
FT : : : : :
1112 GCTACAAAGCCAGAACACCCATTGGACACACTATAGAAACCTACATT 1161
FT : : : : :
768 ysAspIleThrProLeuAlaHisAspSerGluAspTyrLysLeuIleGlu 784
PN : : : : :
1162 GTGCCTTGGCCCTTGCATGCAAGTACGAGTTCAAGATGATTTC 1211
FT : : : : :
785 GlnTyrLeuLeuAsnThrHisAlaProThrHisLysAspTrpSerLeuG1 801
XX : : : : :
1212 CAGTACCTACAATCTACCATGCTCCACACACAGCAGCTATACCATGAC 1261
FT : : : : :
801 uLeuGluGluValPheSerLeuAspArgAspGlyGluLeuAsnLysTyrS 818
XX : : : : :
1262 CTTCGCTGGATTTGTTGAAGTGGAGAGATGGTGAG.....AAAG 1302
XX : : : : :
818 erArgTyrLysAsnAsnLeuHisAsnLysMetLeuLeuTrpHisGlySer 834
XX : : : : :
1303 AAGCCTTACAGAGGACCTTCAACAGAGATGCTTCTATGGCATGGTTCC 1352
XX : : : : :
835 ArgLeuThrAsnPheValGlyIleLeuSerGlnGlyLeuArgIleAlaPr 851
XX : : : : :
1353 AGGATGAGTAAGTGGTGGGAATCTTGAGCCATGGGCTTCAATTTGCCCC 1402
XX : : : : :
851 oProGluAlaProValThrGlyTyrMetPheGlyLysGlyLeuTyrPheA 868
XX : : : : :
1403 ACCTGAAGCTCCATCACAGGTTACATGTTTGGGAAGGAATCTACTTTG 1452
XX : : : : :
868 laAspLeuValSerLysSerAlaGlnTyrCysTyrValAspArgAsnAsn 884
XX : : : : :
1453 CTGACATGCTCTTCAAGAGTGGCAATCTACTGCTTTCCTCGCTCAAG 1502
XX : : : : :
885 ProValGlyLeuMetLeuLeuSerGluValAlaLeuGlyAspMetTyrG1 901
XX : : : : :
1503 AATACAGGACTGCTGCTCTTATCAGAGGTAGCTTAGTTCAGTGAATGA 1552
XX : : : : :
901 uLeuLysLysAlaThr...SerMetAspLysProProArgGlyLysHisS 917
XX : : : : :
1553 ACTACTAGAGGCGCAATCTTAAGGCGGAGGATTTGCTTCAAGGTAACATA 1602
XX : : : : :
917 erThrLysGlyLeuGlyLysThrValProLeuGluSerGluPheValLys 933
XX : : : : :
1603 GCACCAAGGGCTGGGCAAGATGGCTCCCGATTCGCCACCTCGTCACC 1652
XX : : : : :
934 TrpArgAspValValProCysGlyLysProValProSerSerI1 950
XX : : : : :
1653 CTGAATGGGAGT...ACAGTGCCATTAGGACCCAGCAAGTGACACAGGAAT 1699
XX : : : : :
950 eArgSerSerGlu.....LeuMetTyrAsnGluTyrIleValTyrA 964
XX : : : : :
1700 TCTGAATCCAGATGGTTATACCTCAACTACATGAATGAATATATGTATATA 1749
XX : : : : :
964 snThrSerGlnLysMetGlnPheLeuLeuLysValArgPheHis 979
XX : : : : :
1750 ACCCAACCCAGGTCGGTATGCGGTACCTTTAAAGGTTTCAGTTAAT 1796
XX : : : : :
seq_name: /SID52/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ44287

ischemic tissue damage; ss.
Homo sapiens.
Key Location/Qualifiers
CDS 3..1715
/*tag= a
/product= "PARP2"
WO9964572-A2.
16-DEC-1999.
04-JUN-1999; 99WO-EP03889.
05-JUN-1998; 98DE-1025213.
01-MAR-1999; 99DE-1008837.
(BADI) BASF AG.
Kock M, Hoeger T, Kroeger B, Otterbach B, Lubisch W, Lemaire H;
WPI; 2000-087218/07.
P-PSDB; AAY51174.
Novel genes and proteins, antibodies and binding partners useful in
diagnosis and therapy of energy deficiency associated disease
conditions
Claim 7a; Page 49-52; 96pp; German.
This invention describes novel human and murine poly(ADP-ribose)
polymerase (PARP) homologues, which are characterised by an amino acid
sequence with a functional NAD⁺-binding site and no zinc finger
sequence motif, of general formula CX₂CX₂MX₂2C (1). The nucleic acid
sequences, PARP homologues and antibodies are useful for analytic
detection of PARP homologues and for identifying PARP effectors or
binding partners, as well as for determining their effectiveness.
PARP-binding partners are useful for the diagnosis or therapy of a
disease condition, which is the result of a PARP protein, especially an
energy deficiency, which may comprise tissue damage from cell death
following necrosis or apoptosis, the disease condition may be chosen
from a neurodegenerative illness, or sepsis or ischemic tissue damage,
in particular neurotoxic disturbances, etc. This sequence encodes the
human PARP2 protein used in the method of the invention.
Sequence 1843 BP; 599 A; 377 C; 447 G; 420 T; 0 other;

alignment_scores:
Quality: 851.50 Length: 499
Ratio: 2.534 Gaps: 15
Percent Similarity: 67.335 Percent Identity: 40.882
alignment_block:
US-09-236-995D-2 x AAZ44287 ..
Align seg 1/1 to: AAZ44287 from: 1 to: 1843
519 GlnCys***LeuLysHisValLeuThr***His***ValCysThrGly.. 534
: : : : :
252 GAGTGTACAGCCAAG...GTGGGAAGGCTCATCTGTATTGTGAAGGAA 298
: : : : :
534 534
299 TGATGCTCATGATGTCATGCTAAATCAGACCAATCTCCAGTTCAACAACA 348
535TyrTyrValLeuGlnIleLeuGlnAspAspGlySerGlyCys 549
: : : : :
349 ACAAGTACTACTGATTCAGCTATTAGAGATGATGATGCCAGAGAACTTC 398
: : : : :
550 TyrValPheArgLysTrpGlyArgValGlySerGluLysIleGlyGlyG1 566
: : : : :
seq_documentation_block:
ID AAZ44287 standard; cDNA; 1843 BP.
XX
AC
AAZ44287;
DT
31-MAR-2000 (first entry)
XX
XX
DE
Human brain PARP2 cDNA.
XX
KW
PARP; poly(ADP-ribose) polymerase; human; murine; detection; sepsis;
diagnosis; therapy; necrosis; apoptosis; neurodegenerative illness;
KW

851 oProGluLaProValThrGlyTyrMetPheGlyLysGlyLeuTyrPheA 868
1304 ACCTGAAGCTCCACAGGTTCATGTTTGGGAAAGGAATCTCATTTC
868 IaAspLeuValSerLysSerAlaGlnTyrCystTyrValAspArgAsnAsn 884
1354 CTGCACATGCTTCCCAAGAGTGCCTTACGTGCTTGCTCTCGCTAAAG 1403
885 ProValGlyLeuMetLeuSerGluValAlaLeuGlyAspMetTyrGl 901
1404 AATACAGCAGCTGCTCTTATCAGAGTAGCTCTAGGTCAGTGTAAATGA 1453
901 uLeuLysLysAlaThr...SerMetAspLysProProArgGlyLysHis 917
1454 ACTACTAGAGGCCAATCCTTAAGGCCGAAGGATGCTTCAAGGTAAACATA 1503
917 erThrLysGlyLeuGlyLysThrValProLeuGluSerGluPheValLys 933
1504 GCACCAAGGGCTGGCAAGATGGCTCCAGTCTCTGCCACTTCGTCAACC 1553
934 TrpArgAspAspValValProCysGlyLysProValProSerSerII 950
1554 CTGAATGGAGT...ACAGTGCCTTATAGCACCAAGTGCACAGGAAT 1600
950 eArgSerSerGlu.....LeuMetTyrAsnGluTyrIleValTyrA 964
1601 TCTGAATCCAGATGGTTATACCCCTCACTACATGAATATATGTTATATA 1650
964 snThrSerGlnValLysMetGlnPheLeuLeuLysValArgPheHis 979
1651 ACCCCACCAAGGTCGGTATGCGGTACCTTTTAAAGGTTTCAGTTTAAT 1697

seq_documentation_block:

XX			
AC	AAC82090;		
XX			
DT	02-MAR-2001	(first entry)	
XX			
DE	Human brain	poly-ADP-ribose-polymerase	cDNA

OS Homo sapiens.

PD 16-NOV-2000.

PR 11-MAY-1999; 99DE-1021567.

PI Lubisch W, Sadowski J, Kock M, Hoeger T;

Drugs for inhibiting PARP or especially homologous enzymes comprising 4-substituted phthalazine derivatives, useful e.g. for treating neurodegenerative disease, ischemic damage, tumors or diabetes -

XX Example A; Page 9-12; 14pp; German.

XX This invention describes novel 4-substituted 2H-phthalazin-1-one

XX derivatives (I) which are used for the treatment or prophylaxis of

XX diseases associated with elevated poly-(ADP-ribose)-polymerase (PARP;

XX EC 2.4.2.30) activity. The products of the invention have neurotropic,

XX neuroprotective, cerebroprotective, antiparkinsonian, nephrotropic,

XX cardiant, vasotropic, anticonvulsant, cytostatic, antibacterial,

XX immunosuppressive, antiinflammatory, antirheumatic, antiarthritic,

XX antidiabetic. (I) are especially used for treating or preventing

XX neurodegenerative disease or neuronal damage (specifically associated

XX with ischemia, trauma or massive bleeding, especially apoplexy or

XX spinal-cranial trauma; or Alzheimer's disease, Huntington's disease or

XX Parkinson's disease), treating or preventing ischemic damage

XX (specifically renal damage after renal ischemia or during and after

XX kidney transplantation or heart damage after cardiac ischemia), treating

XX epilepsy, specifically generalized epileptic attacks (e.g. petit mal and

XX tonic-clonic attacks) or partial epileptic attacks (e.g. temporal lobe

XX and complex partial attacks), treating microinfarction (e.g. during and

XX after heart valve replacement, aneurysm resectioning and heart

XX transplantation), revascularization of critically constricted coronary

XX arteries (e.g. after PCTA or by-pass operations) or peripheral arteries

XX (e.g. leg arteries), treating acute myocardial ischemia and damage during

XX or after its mechanical or drug-induced lysis and treating tumors and

XX their metastasis, sepsis and septic shock, inflammatory and rheumatic

XX disease (e.g. rheumatoid arthritis), and diabetes mellitus. Although (I)

XX inhibit PARP (i.e. the known form designated PARP1), they especially

XX selectively and strongly inhibit PARP homologs, specifically the homolog

XX PARP2. In particular (I) have very strong PARP2 inhibitory activity (e.g.

XX with K_i values of 1-20 nM) and high selectivity for PARP2 relative to

XX PARP1 (generally by a factor of more than 5).

SQ Sequence 1843 BP; 599 A; 377 C; 447 G; 420 T; 0 other;

alignment_scores:

Quality: 851.50 Length: 499

Ratio: 2.534 Gaps: 15

Percent Similarity: 67.335 Percent Identity: 40.882

alignment_block:

US-09-236-995D-2 x AAC82090 ..

Align seg 1/1 to: AAC82090 from: 1 to: 1843

519 GlnCys***LeuLysHisValLeuThr***His***ValCysThrGly... 534

252 GAGTGTACAGCCAAAG...GTGGGGAAGGCTCATGTATTGTGAAGGAAA 298

534 534

299 TGATGTCTATGATGCTAATCAACAGCAATCTCCAGTTCAACACA 348

535TyrTyrValLeuGlnIleLeuGlnAspGlySerGluCys 549

349 ACAAGTACTATCTGATTCAGCTATTAGAGATGATGCCAGAGGAATTC 398

550 TyrValPheArgLysTrpGlyArgValGlySerGluLysIleGlyGlyG 566

399 AGTGTGTTGGATGAGATGGGGCCGAGTTGGG.....AAAATGGGACAGCA 442

566 nLysLeuGluGluMetSer.....LysThrGluAlaIleLysGluPheL 581

443 CAGCCTGTGGTGTGTTGTCAGCAATCTCAACAAAGGCCAAGGAAATCTTTC 492

581 ysArgLeuPheLeuGluLysThrGlyAsnSerTrpGluAlaTrpGluCys 597

493 AGAAGAAATTCCTTGACAAAACGAAAAACAATGGGAGATCGACAAA... 539

598 LysThrAsnPheArgLysGlnProGlyArgPheTyrProLeuAspValAs 614

540AAGTTTCAGAAAGGTGCTCGGAAATATGATATGATACAGATGA 583

614 pTyrGlyValLysLysAlaProLysArgLysAspIleSerGlu.....M 629

584 CTATGCCACCAATACTCAGCATGAAGGAAACAAAGAAAGAGATCTC 633

629 eLysSerSerLeuAlaProGln.....LeuLeuGlu 639

634 TTAATCTCCCTTGAAGCCACAGACTCACAGCTAGATCTTCGGGTACAGGAG 683

640 LeuMetLysMetLeuPheAsnValGluThrTyrArgAlaAlaMetMetG 656

684 TTAATAAAGTTGATCTGTATGTCAGGCCATGGAAGAAATGATGATGGA 733

656 uPheGlu***AsnMetSerGluMetProLeuGlyLysLeuSer*****A 673

734 AATGAAGTATTAATACCAAGAAAGCCCACTTGGGAAGCTGACAGTGCAC 783

673 snIleGlu***GlyPheGluAlaLeuThr*****LeuPheGlu 689

784 AATCAAGGCAAGGTTACCACTCTCTTAAGAAGATTCAGGATTCGTATTCGG 833

690 GlyHisArg***SerSerThrGlyLeu***GluLysAla***LeuLeu** 706

834GTTGGCCAGCATGAGCAGCTCTCATGGAAGC 865

706 *****PheSerLeuLeuSerLeuLeuPheIleLeu..... 720

866 ATGCAATGAATTCACACCAAGGATTCGGCATGACTTTGGACTCCGTACTC 915

721IleLeuTyrGlyMetArgMetIleSerTyrSerLysAlaLysMet 735

916 CTCACCTAATCCGACACACAGAGAACTGTCA...GAAATAATACAAATTA 962

736 LeuGluAlaLeuGlnAspIleGluIleAlaSerLysIleValGlyPheAs 752

963 CTAGAGCCTTTGGGAGACATTGAAATTCCTATTAAAGCTGGTGAAACAGA 1012

752 pSerAspSerAspGlu...SerLeuAspAspLysTyrMetLysLeuHisC 768

1013 GCTACAAGCCACAGACACCCATTTGGACCAACACTATAGAAACCTACATT 1062

768 ysAspIleThrProLeuAlaHisAspSerGluAspTyrLysLeuIleGlu 784

1063 GTGCCTTTCGCCCTTGCACCTGACCAAGATTACGAGTTCAAGATGATTTCC 1112

785 GlnTyrLeuLeuAsnThrHisAlaProThrHisLysAspTrpSerLeuG 801

1113 CAGTACCTACAATCTACCCATGCTCCACACACAGCAGCACTATACCATGAC 1162

801 uLeuGluGluValPheSerLeuAspArgAspGlyGluLeuAsnLysTyrS 818

1163 CTGCTGGATTTGTTGAAGTGGAGGATGGTGAG.....AAG 1203

818 erArgTyrLysAsnAsnLeuHisAsnLysMetLeuLeuTrpHisGlySer 834

1204 AAGCCTTCAGAGAGGACCTTCATAACAGGATGCTTCTATGGCATGGTTCC 1253

835 ArgLeuThrAsnPheValGlyIleLeuSerGlnGlyLeuArgIleAlaPr 851

1254 AGGATGAGTAACCTGGGGGGAATCTTGAGCCATGGGCTTCGAATTCGCC 1303

851 oProGluAlaProValThrGlyTyrMetPheGlyLysGlyLeuTyrPheA 868

1304 ACCTGAGCTCCCATCACAGGTTCATGTTTGGGAAGGAATCTACTTTG 1353

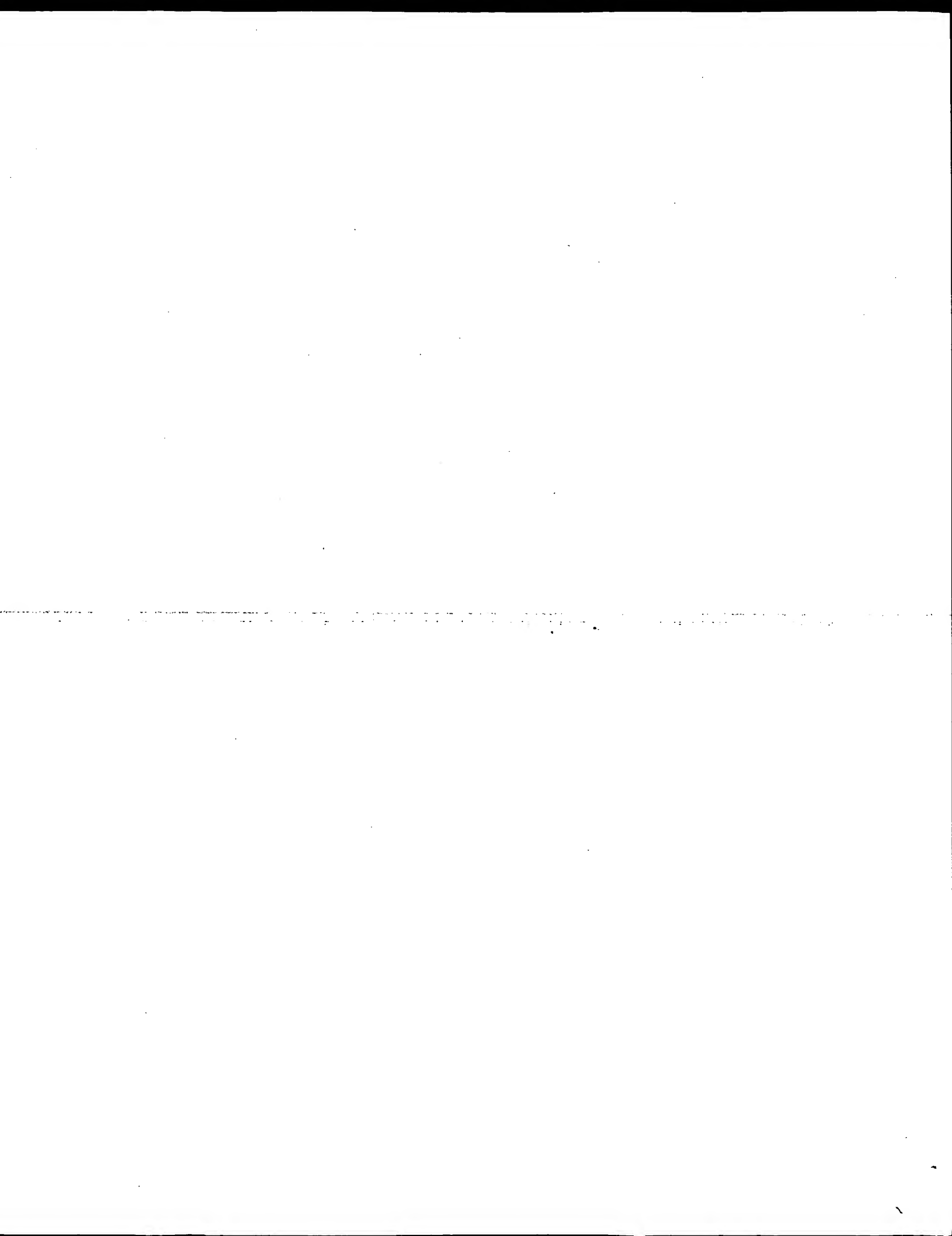
868 laAspLeuValSerLysSerAlaGlnTyrCysTyrValAspArgAsnAsn 884

1354 CTGACATGTCTTCCAGAGAGTGCCTTACCTGCTTTGCTCTCGCCTAAAG 1403

885 ProValGlyLeuMetLeuLeuSerGluValAlaLeuGlyAspMetTyrG 901

1404 AATACAGGAGTGTGCTCTTATCAGAGGTAGCTCTAGGTCAGTGTATGA 1453

950 eARGSerSerGlu.....LeuMetTyrAsnGluTyrIleValTyrA 964
 1655 TCTGAATCCAGATGGTATACCCCTCAACTACAAAGATATTTGTATATA 1714
 964 snThrSerGlnValLysMetGlnPheLeuLeuLysValArgPheHis 979
 1715 ACCCAACACCGATCGGTATGGGTACCTTTTAAAGGTTCAGTTAAT 1761



Mon Mar 11 09:58:42 2002

309 AGTCGCCCATGTTTGGTGAAGGATCCACACTGGTACCACCTTCTCTGCG 358
56 IlePheSerLysLysAsnGlnIleLysSerValAsp...AspValGluG1 71
359 TTCTCGAAGGTGGCCCATCTCCATCGGACCCCTGACGTTGAGGTGGATGG 408
71 yIleAspAlaLeuArgTrrpAspAspGlnGluLysIleArgAsnTyrVal. 87
409 GTTCTCTGAGCTTCGGTGGGTGACACAGAGTCAAGAGACAGCGG 458
88GlySerAlaSerAlaGlyThrSerSerThrAlaAlaProPro 101
459 AAGCTGGAGGATGACAGCAAGCCAGGATGGAAATGGTAGCAAGCA 508
102 GluLysCys.....ThrIleGluIleAlaProSerAlaArgTh 114
509 GAGAGACTCTGGGTGACTTTCGACGAGAGTATGCCAAGTCCACAGAG 558
114 rSerCysArgArgCysSerGluLysIleThrLysGlySerValArgLeuS 131
559 TAGCTCAAGGGGTGTATGGAGAAGATAGAAAAGGCCAGGTGCGCCTGT 608
131 eAlaLysLeu...GluSerGluGlyProLys...GlyIle.....Pro 143
609 CCAAGAGATGGTGGACCGGAGAGCCACACTAGGATGATTGACCGC 658
144 TrpTyrHisAlaAsnCysPhePheGlu.....ValSe 154
659 TGCTACCATCCAGGCTGCTTTGTCAAGAACAGGAGGAGCTGGGTTCG 708
154 rProSerAlaThrValGluLysPheSerGlyTrpAspThrLeuSerAspG 171
709 GCCCGAGTACAGTGCAGTGCAGTCAAGGGCTTCAGCCTCCTTGCTAC 758
171 luAspLysArgThrMetLeuAspLeuValLysLysAspValGlyAsnAsn 187
759 AGGATAA.....GAAGCCCTGAAGAAGCAGCTCCAGGAGTC 796
188 GluGlnAsnLysGlySerLysArgLysLysSerGluAsnAspIleAspSe 204
797 AAG.....AGTGAAGAAAGAGAAAGAGGCGATGAGGTGGATGG 834
204 rTyrLysSerAlaAlaLeuAspGluSerThrSerGluGlyThrValArgA 221
835 A.....GTGGATGAGTGGCGAAGAGAAATCTAATAAAG 869
221 snLysGlyGlnLeuValAspProArgGlySerAsnThrSerSerAlaAsp 237
870 AAAAGACAAG.....GATAGTAAG 889
238 IleGlnLeuLysLysGluGlnSerAspThrLeuTrpLysLysAs 254
890 CTTGAAAAGCCCTAAAGGCTCAGACGACCTGATCTGGAACATCAAGGA 939
254 pGlyLeuLysThrHisValSerAlaAlaGluLeuArgAspMetLeuGluA 271
940 CGAGCTAAAGAAAGTGTGTCAACTAATGACCTGAAGGAGCTACTCATCT 989
271 laAsnGlyGlnAspThrSerGlyProGluArgHisLeuLeuAspArgCys 287
990 TCAACAAGCAGCAAGTCCCTTGGGAGGTCGGGATCTTGGACCGAGTA 1039
288 AlaAspGlyMetLeuPheGlyAlaLeuGlyProCysProValCysAlaAs 304
1040 CTTGATGGCATGGTTCGGTGGCCCTCTTCCCTGCGAGGAATGCTCGGG 1089
304 nGlyMetTyrTyrAsnGlyGlnThrGlnCysSerGlyAsnValSerG 321
1090 TCAGCTGGTCTTCAAGACGAGTGGCTATTACTGCACCTGGGAGCTCACTG 1139
321 luTrpSerLysCysThrTyrSerAlaThrGluProValArgValLysLys 337
1140 CTGGACCAAGTGTATGCTCAAGACACAGACACCAACCGG.....AAG 1183

338 LysTrpGlnIleProHisGlyThrLysAsnAspTyrLeuMetLysTrpPh 354
1184 GAGTGGGTAAACCCA.....AAGAAATCCGAGAAATCTCTTACCT 1224
354 eLysSerGlnLysValLysLysProGluArgValLeuProProMetSer. 370
1225 CAAGAAATGAAGTTAAAAAGCAGGACCGTATATTCCTCCCAAGAACCA 1274
371ProGluLysSerGlySerLysAla 378
1275 CGCCTCCGTGGCGGCACCCCTCCCGCTCCACAGCCTCGGCTCTCTGCT 1324
379 ThrGlnArgThrSerLeuSerSerLysGlyLeuAspLysLeuArgPh 395
1325 GCTGTGAACCTCCTCTGCTCAGCAGATAAGCCATTATCCAAACATGAAGAT 1374
395 eSerValValGly.....GlnSerLysGluAlaAlaAsnGluTrpI 409
1375 CCTGACTCTCGGAAGCTGTCCCGGAACAAGATGAAGTGAAGGCCCATGA 1424
409 leGluLysLeu.....LysLeuAlaGlyAlaAsnPheTyrAlaArgVal 423
1425 TTGAGAAACTCGGGGGAAGTTGACGGGAGCGGCCAACACAGGCTTCCCTG 1474
424 VallysAspIleAspCysLeuIleAlaCysGlyGluLeuAspAsnGluAs 440
1475TGCATCAGCACCAACAAAGAGGAGGTGGAAGAGATGAA 1509
440 nAlaGluValArgLysAlaArgArgLeuLysIleProIleValArgGluG 457
1510 TAAGAAGATGGAGGAAGTAAAGGAAGCCACATCCGAGTGTGTCTGAGG 1559
457 LyTyrIleGlyCysValLysArgThrLys..... 467
1560 ACTTCTCCAGCAGCTCTCCCTCCACCAAGAGCCTTCAGGAGTGTGTC 1609
468 CysCysHisLeuIleCysIleAsnTrpAsnAlaLeuGluSerLysG1 484
1610 TTAGCGCACATCTTGTCC...CCTTGGGGGCGAGAGGTGAAGGCAGAGCC 1656
484 y**ThrValThrValLysValLysGlyArgSerAlaCysSer***SerP 501
1657 TGTGAAGTT...GTGGCCCCAAGAGGAACTCAGGGGCTCGCTCTCCA 1703
501 ro**ValCysLysAsnThrAlaHisIlePro*****Trp..... 514
1704 AA.....AAAAGCAAGGCCAGGTCAAGAGGAGGATATCAACAAA 1744
515 ...GluLysHisIleGlnCys***LeuLys..... 523
1745 TCTGAAAAGAGAAATGAAATTAATCTTTAAAGGAGGAGCAGCTGTGTGATCC 1794
524HisValLeu..... 526
1795 TGATTTGGACTGGAACACTCTCGCGCATGTCTCTGGAGAAAGGTGGGAAGG 1844
527Thr***His***Val.....CysThrGly 534
1845 TCTTCAAGTCCACCTTGGCCTGGTGGAGACATGCTTAAAGGAACCACTCC 1894
535 TyrTyrValLeuGlnIleIleGluGlnAspAspGlySerGluCysTyrVa 551
1895 TACTACAAGCTCGAGCTTCTGAGGAGCAGACAAAGAAACAGGTATGGAT 1944
551 lPheArgLysTrpGlyArgValGlySerGluLysIleGlyGlnLysL 568
1945 ATTCAAGTCTCGGGCGGTGTGGGTACG...GTGATCGGTAGCAACAAAC 1991
568 euGluGluMet...SerLysThrGluAlaIleLysGluPheLysArgLeu 583
1992 TGGACAGATGCGGTCCCAAGGAGGATGCCATTGACAGGTTTCATGAATTA 2041

584 PheLeuGluLysThrGlyAsnSerTrpGluAlaTrpGluCysLysThrAs 600
2042 TATGAAGAAAAACCGGAGCGCTTGGCAGCTCC.....AAAAA 2079
600 nPheArgLysGlnProGlyArgPheTyrProLeuAspValAspTyrGlyV 617
2080 TTTACGAGAGTATCCCAAAAGATTTTACCCCTGGAGATTGACTATGGCC 2129
617 al....LysLysAlaProLysArgLysAspLysSerGlu...MetLysSer 631
2130 AGGATGAAGAGCGAGTGAAGAAGCTCAGCTAAATCTCGGCCCAAGCTCC 2179
632 SerLeuAlaProGlnLeuLeuGluLeuMetLysMetLysMetLysPheAsnValGI 648
2180 AAGCTCCCAAGCCAGTTCAGGACCTCATCAAGATGATCTTTGATGTGA 2229
648 uThrTyrArgAlaAlaMetMetGluPheGlu***AsnMetSerGluMetp 665
2230 AAGTATGAAGAAACCCATGGTGGAGTATGAGATCGACCTTCAGAAAGATGC 2279
665 roLeuGlyLysLeuSer*****AsnIleGlu***GlyPheGluAlaLeu 681
2280 CCTTGGGAGCTGAGCAAAAGGAGATCCAGGCCGATATCTCCATCCCTC 2329
682 Thr*****LeuPheGluGlyHisArg***SerSerThrGlyLe 698
2330 AGT.....GAGGTCCAGCAGCGGTGTCTCAGGCGAG 2361
698 u**GluLysAla***LeuLeu*****PheSerLeuLeu 715
2362 CAGGACTCTCAGATCTCGATCTCTCAATCGCTTTTACACCTGATCC 2411
715 erLeuLeuPheIleLeuIleLeuTyrGlyMetArgMet..... 727
2412 CCCACGACTTT.....GGGATGAAGAAGCTCCGCTCCCTG 2446
728IleSerTyrSerLysAlaLysMetLeuCluAlaLeuGI 740
2447 AACATTCACAGAGTGTGACGAGCGAGGTGAAATCTTGCACACCTGCT 2496
740 nAspIleGluIleAlaSerLysIleVal.....GlyPheAspSerAsps 755
2497 GGACATCGAGGTGCTTACAGTCTGCTCAGGCGAGGTCTGATGATAGCA 2546
755 erAspGluSerLeuAspLysTyrMetLysLeuHisCysAspIleThr 771
2547 GCAAGATCCCATCGATGTCAACTATGAGAAGCTCAAACTGACATTAAG 2596
772 ProLeuAlaHisAspSerGluAspTyrLysLeuIleGluClnTyrLeuLe 788
2597 GTGTTGACAGAGATTCTGAAGAAGCCGAGATCATCAGGAAGTATGTAA 2646
788 uAsnThrHisAlaProThrHisLysAspTrpSerLeuGluLeuGluV 805
2647 GAACACTCATGCAACACACACAGTCCGTATGACTTGGAGTCATCGATA 2696
805 alPheSerLeuAspArgAspGlyGluLeuAsnLysTyrSerArgTyrLys 821
2697 TCTTTAAGATAGCGGTGAAGCGAATGCGAGCGTTACAGCCCTTTAAG 2746
822 AsnAsnLeuHisAsnLysMetLeuLeuTrpHisGlySerArgLeuThrAs 838
2747 ...CAGCTTCATAACCGAAGATTGCTGGCAGCGGTCCAGGACCAACAA 2793
838 nPheValGlyIleLeuSerGlnGlyLeuArgIleAlaProGluAlaP 855
2794 CTTTGTGGATCCTCTCCAGGGTCTTCGGATAGCCCGCCCTGAAGCGC 2843
855 roValThrGlyTyrMetPheGlyLysGlyLeuTyrPheAlaAspLeuVal 871
2844 CCGTGACAGGCTACATGTTTGTGAAGGATCTATTTCGCTGACATGTC 2893
872 SerLysSerAlaGlnTyrCysTyrValArgAsnAsnProValGlyLe 888

2894 TCCAAGAGTGCCAACTACTACCATAGTCTCAGGAGACCAATAGGCTT 2943
888 uMetLeuLeuSerGluValAlaLeuGlyAspMetTyrGluLeuLysLysA 905
2944 AATCCCTGTTGGGAAAGTTCCCTTGGAAACATGTATGAAGTGAAGCAG 2993
905 laThrSerMetAspLysProProArgGlyLysHisSerThrLysGlyLeu 921
2994 CTTACATATCAGCAGGTTACCCAGGAGGACACAGTGTCAAGGTTTG 3043
922 GlyLysThrValProLeuGluSerGluPheValLysTrpArgAspVa 938
3044 GCAAAATACCCCTGATCTCTCAGCTAACATTAGT...CTGGATGGTCT 3090
938 lValValProCysGlyLysProValProSerSerIleArgSerSerGluL 955
3091 AGAGTTCCTCTTGGGACCGGATTCATCTGCTGTAATGACACCTCTC 3140
955 euMetTyrAsnGluTyrIleValTyrAsnThrSerGlnValLysMetGln 971
3141 TACTATATAACAGTACATGTCTATGATATTGCTCAGGTAATCTGAAG 3190
972 PheLeuLeuLysValArgPheHisLysLys 981
3191 TATCTGCTGAAACTGAAATCAATTTAAG 3220

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-044-618-7

seq_documentation_block:

; Sequence 7, Application US/08044618

; Patent No. 5449605

; GENERAL INFORMATION:

; APPLICANT: SMULSON, MARK

; TITLE OF INVENTION: METHOD OR DETECTING A PREDISPOSITION TO

; TITLE OF INVENTION: CANCER BY THE USED OF RESTRICTION FRAGMENT LENGTH

; TITLE OF INVENTION: POLYMORPHISM OF THE GENE FOR THE HUMAN POLY (ADP-RIBOSE)

; NUMBER OF SEQUENCES: 7

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Sterne, Kessler, Goldstein & Fox

; STREET: 1225 Connecticut Suite 300

; CITY: Washington

; STATE: D.C.

; ZIP: 20036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/044,618

; FILING DATE: 19930406

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/07/257,696

; FILING DATE: 14-OCT-1988

; ATTORNEY/AGENT INFORMATION:

; NAME: FOX, SAMUEL L

; REGISTRATION NUMBER: 30,353

; REFERENCE/DOCKET NUMBER: 0654.0490001

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202)466-0800

; TELEFAX: (202)833-8716

; INFORMATION FOR SEQ ID NO: 7:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 5345 base pairs

; TYPE: NUCLEIC ACID

; STRANDEDNESS: both

; TOPOLOGY: linear

; MOLECULE TYPE: DNA

; US-08-044-618-7

[illegible][illegible]

506 snThrAlaHisIlePro*****Trp..... 514
||||| : : : : : |||||
2385 ACACCTGCGCATG.....TCCTGGAGAAAGGTGAAGGCAGAGCCTGTT 2428
514 514
2429 GAAGTCGTAGCCCCAAGAGGAACTCAGGAGCTGTGCTCTCCAAAAAAG 2478
515 GluLysHisIleG 519
2479 CAAGGGCCAGGTCAAGGAGGAGGTATCAACAAATCTGAAACAGAAATGA 2528
519 InCys***LeuLys..... 523
2529 AATTAACCTCTTAAAGGAGGAGGAGCAGCTGTGATCTGCTGCTGGAA 2578
524HisValLeu.....Thr** 528
2579 CACTCTGCACATGTCGTGGAGAAAGGTGGGAAGGCTTTCAGTCCACCCCT 2628
528 *His***Val.....CysThrGlyTyrValLeuGlnI 540
||| : : : : : ||||| : : : : :
2629 CAGCCTGTGTCGCTGTTAAAGGAACCACTCTCTATTAACAGCTGAAGT 2678
540 leileGluInAspAspGlySerGluCysTyrValPheArgLysTrpGly 556
: : : : : ||||| : : : : : ||||| : : : : :
2679 TGCTGAAGGATGACAGGAAGGAGGAGCACTTGGATATTAAGTCTGGGAC 2728
557 ArgValGlySerGluLysIleGlyGlnLysLeuGluMet...Se 572
||||| : : : : : ||||| : : : : : ||||| : : : : :
2729 CGTGTGGGACG...GTGATGCGGTAGCAACAACTGGAACAGATGCTGTG 2775
572 rLysThrGluAlaIleLysGluPheLysArgLeuPheLeuGluLysThrG 589
||||| : : : : : ||||| : : : : : ||||| : : : : :
2776 CAAGGAGGACACCATTCACACTTCATGAAATATATGAAGAAACTAG 2825
589 lyAsnSerTrpGluAlaTrpGluCysLysThrAsnPheArgLysGlnPro 605
|| : : : : : ||||| : : : : : ||||| : : : : :
2826 GAAT.....GCTTGGCAGCTCCAAAAA...TTCACAAAGTATCCC 2861
606 GlyArgPheTyrProLeuAspValAspTyrGlyVal...LysLysAlaPr 621
: : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
2862 AAAAAAGTCTACCCCTGGAGATGACTACGCCAGGACAAAGAGCGGT 2911
621 oLysArgLysAspIleSerGluMetLysSerSerLeuAlaProGln...L 637
: : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
2912 GAAGAGCCGACAGTAATCCTGGCACCACCATGCTCTCCCAAGCCAG 2961
637 euLeuGluLeuMetLysMetLeuPheAsnValGluThrTyrArgAlaAla 653
: : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
2962 TTCAGGACCTCATCAAGATGATCTTTGATGTGGAAGATGAGCAAGCC 3011
654 MetMetGluPheGlu***AsnMetSerGluMetProLeuGlyLysLeuSe 670
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
3012 ATGCTGGGTGTGAGATCAACCTT...CAGATGCCCTTGGGAAGCTGAG 3058
670 r*****AsnIleGlu***GlyPheGluAlaLeuThr*****L 687
|| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
3059 CAAAGGCAAAATCCAGGCGCGTACTCTCCATCTCAGGTC..... 3097
687 euPheGluGlyHisArg***SerSerThrGlyLeu***GluLysAla*** 703
: : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
3098CAGCAGGTGGTGTCCAGGGGAGGAGGAGCACTCTCAGATC 3136
704 LeuLeu*****PheSerLeuLeuSerLeuPheIleLe 720
||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
3137 CAG.....GATCTCTCAAAATCGCTTTTACATCT 3165
720 uLeuLeu.....TyrGlyMetArg.....MetIleSer..... 729
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
3166 GATCCCCCAGCAGCTTGGGATGAAGGATCTCTGCTCTGCTGGAACATGCGAG 3215

730TyrSerLysAlaLysMetLeuGluAlaLeuGlnAspIleGlu 743
: : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
3216 ACAGTGTGCGAGCCCAAGGTAGAATCTGGCAACACTGCTGGACATGAG 3265
744 IleAlaSerLysIleVal.....GlyPheAspSerAspSerAspGluSe 758
: : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
3266 GTAGCCTACGGTCTGCTCAGGGGAGGCTTCCAGATAGCAGGAAGGACTC 3315
758 rLeuAspAspLysTyrMetLysLeuHisCysAspIleThrProLeuAlaH 775
||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
3316 CATCATGTCAACTATGACAAGCTCAAAACTCACAATTAAAGGTGGTTGACA 3365
775 isAspSerGluAspTyrLysLeuIleGluGlnTyrLeuLeuAsnThrHis 791
: : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
3366 GAGATCTCTGAAGAGCTGAGATCATCAGGAAGTATGTTAAGAACACTCAT 3415
792 Ala.Pro...ThrHisLysAspTyrSerLeuGluLeuGluValPheS 807
||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
3416 GCAACCAACCAACACAGATGATATGACTTGGAGTCAATGATAGCTTTTA 3465
807 erLeuAspArgAspGlyGluLeuAsnLysTyrSerArgTyrLysAsnAsn 823
: : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
3466 AGATAGAGTGTGAAGAGGAGTCCAGCACTACAGCCCTTTAAG...CAG 3512
824 LeuHisAsnLysMetLeuLeuThrPheHisGlySerArgLeuThrAsnPheVa 840
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
3513 CTTCACTAACTGAAGCTTCTGTGGCATGGTCCAGGACCACCACTTGC 3562
840 lGlyIleLeuSerGlnGlyLeuArgIleAlaProProGluAlaProValT 857
: : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
3563 TGGGATCTGTCCTCCCTGGGTCTTGGATAGCCCTGCTGGAACCACTGTGA 3612
857 hrGlyTyrMetPheGlyLysGlyLeuTyrPheAlaAspLeuValSerLys 873
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
3613 TGGGTACATGTTGGTAAAGTATGATCTATTCGCTGATCTTGTCTCCAAG 3662
874 SerAlaGlnTyrCysTyrValAspArgAsnAsnProValGlyLeuMetLe 890
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
3663 AGTGCACAGCACTGCCATACATCTTAGGAAGACCAATAGGTTAATCCT 3712
890 uLeuSerGluValAlaLeuGlyAspMetTyrGluLeuLysLysAlaThrS 907
|| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
3713 GTCGGAAGAAGTGGCCTTGGAAAGCTGTGGAACCTGAAGCATGCTTCA 3762
907 erMetAspLysProArgGlyLysHisSerThrLysGlyLysGlyLys 923
: : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
3763 ATATCAGCAAGTTACCAAGGCAAGCACTGTCAAGGTTTGGGCAAA 3812
924 ThrValProLeuGluSerGluPheValLysTrpArgAspValValva 940
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
3813 ACTACTCTGACCTTTCAGCTAGTATCCCA...CTGGATGGTGTAGAGT 3859
940 lProCysGlyLysProValProSerSerIleArgSerSerGluLeuMetT 957
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
3860 TCCTCTTGGGACCGGGTTTCATCTGGTGTGAATGACACCTGTCTACTGT 3909
957 yrAsnGluTyrIleValTyrAsnThrSerGlnValLysMetGlnPheLeu 973
: : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
3910 ATAATGATACATGCTCTATGATATGCTCAGGTAATCTCAAAATATCTG 3959
974 LeuLysValArgPheHisHisLys 981
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
3960 CTGAACCTGAAATTCATTTAAG 3983

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-044-618-3

seq_documentation_block:

; Sequence 3, Application US/08044618

; Patent No. 5449605

; GENERAL INFORMATION:

; APPLICANT: SMULSON, MARK

; TITLE OF INVENTION: METHOD OR DETECTING A PREDISPOSITION TO

; TITLE OF INVENTION: CANCER BY THE USED OF RESTRICTION FRAGMENT LENGTH

298 GAATAGGTGATCCCA.....CGAGAAATCTCTTAACCTCAAGAAT 338
354 helysSerGlnLysValLysLysProGluArgValLeuPro..... 368
339 TT.....AAGGTTAAAGAGCAGGACCGTATATTTCCCGCAAGAAC 379
369MetSerProGluLysSerGlySerLys...Al 378
380 AGTCACCCAGTGGCGCCATGCCCCACCCCTCCACAGGCTCGAGCGCTGC 429
378 aThrGlnArgThrSerLeuLeuSerSerLysGlyLeuAspLysLeuArgP 395
430 TGCATGAACGCTCTGCTCCACAGATAGCGGTATTCACACATGAAGA 479
395 heSerValValGly.....GlnSerLysGluAlaAlaAsnGluTrp 408
480 TCCTTACTCTTGGAGAGCTCTCCGGGAACAAGATGAAGTGAAGGCCACG 529
409 IleGluLysLeuLysLeuAlaGlyAlaAsnPheTyrAlaArgValVal 425
530 ATTGAGAAACTC.....GGGGAAAGTTGATGGGACGCGCAACAA 570
425 sAspIleAspCysLeuLleAlaCysGlyGluLeuAspAsnGluAlaG 442
571 GGCTTCTCCGTCATCAGCACATAAGAGGAGGTGGAAAGAGAGAGAA 620
442 luValArgLysAlaArgArgLeuLysIleProIleValArgGluGlyTyr 458
621 AGATGGAGAGTGAAGAAGCAACATCTGAGCTGTGCTCAGGACTTC 670
459IleGlyGlu.CysValLysA 465
671 CTCAGGACTTCTCTGCTCCACCAAGGGGCTCCAGGAGTTGTTCTCAGC 720
465 rgThr..... 466
721 GCACATATTGACCCCTGGGGGCAGAGGTGAAGGCAGAGCCTGTTGAAG 770
467LysCysCysHis..... 470
771 TCGTAGCCCCAAGAGGAAGTCAGGAGCTGTGCTCTCCAAAAAGAACAG 820
470 470
821 GCCAGGTCAAGGAGGAAGTATCAACAAATCTGAAGAGAGATGAATTT 870
471LeuIleCysIleAsnTrpAsnAla 478
871 AACTCTTAAGAGGAGCAGCTGTGGATCTGACTCTGTGCTGGAACACT 920
479 LeuGluSerSer..... 482
921 CTGCGCATGCTCTGGACAAAGGTGAAGGCAGAGCCTGTTGAAGTCGTAGC 970
483LysGly***ThrValThrValLysValLysGlyArg. 494
971 CCCAAGAGGGAAGTCAGGAGCTGTGCTCTCCAAAAAGCAAGGGCCAGG 1020
494 494
1021 TCAAGGAGGAGGTATCAACAACTGAAAGAGAGATGAATTAACCTCTT 1070
495SerAlaCysSer***SerPro***ValCysLysAsnThrLah 509
1071 AAAGGAGGAGCAGCTGTGGATCTCTGACTCTGTGCTGAACACTCTGACA 1120
509 sIlePro*****TrpGluLysHisIleGlnCys***LeuLysHisVal 526
1121 TGTTCTGGAGAAAGTGGAGAGTCTTCAGTGGCCACCTCAGCCTCGCTGTG 1160
526 euThr***His***ValCysThrGly.....TyrTyrValLeuGln 539

238 IleGlnLeuLysLeuLysGluGlnSerAspThrLeuTrpLysLeuLys 254
: : : : : ||||| : : : : : : : : : : : : : : :
4 CTTGAAGGCCCTAAAGGCCACAGCAGCTGATCTGGAACTCAAGA 53

254 pGlyLeuLysThrHisValSerAlaAlaGluLeuArgAspMetLeuGluA 271
: ||| ||| : : : : : : : : : : : : : : :
54 CGAGCTTAATGAAGTGCTCTATTAAATGACTGAAGGAGATGCTCATCT 103

271 laasnGlyGlnAspThrSerGlyProGluArgHisLeuLeuAspAlaCyS 287
: ||| ||||| : : : : : : : : : : : : : : :
104 TCAACAGGCAGCAGGTGCTTCCGGGGAGTCGGCGATCTTGGACCAGTA 153

288 AlaAspGlyMetLeuPheGlyAlaLeuGlyProCysProValCysAlaAs 304
||||| : : : : : ||||| ||||| ||||| : : : : :
154 GCTGACAGATCATGCTTCGGTGCCTTCTCCCTGTGAGGAATGCTCAGG 203

304 nGlyMetTyrrTyrrAsnGlyGlnTyrrGlnCysSerGlyAsnValSerC 321
: : : : : : : : : : : ||| ||||| : : : : : : : : : :
204 TCAGCTGGTGCTTCAAGACGACACTTATTACTGCACCGGGAGCTCACGG 253

321 luTrpSerLysCysThrTyrrSerAlaThrGluProValArgValLysLys 337
: : : : : ||||| : : : : : ||||| : : : : : : : : : :
254 CTTGGACCAAGGTATGGTCAAGACATGGACACCC.....AACCGAAG 297

338 LvsTrpGlnIleProHisGlyThrLysAsnAspTyrrLeu.MetLysTrpP 354

1170GAGCTGCTTAAAGGAACCAACTCCTCTATTACAGTGAAGT 1208
540 IleIleGluGlnAspAspGlySerGluCysTyrValPheArgLysTrpG1 556
1209 TTGCTGAAGGATGACAGGAAGACAGGCAITGGATATTCAAGTCTGGGA 1258
556 yArgValGlySerGluLysIleGlyGlnLysLeuGluGluMet...S 572
1259 CCGTGTGGGACG...GTGATCGGTAGCAACAACAACTGGAACAGATGCTGT 1305
572 erLysThrGluAlaIleLysLuphLysArgLeuPheLeuGluLysThr 588
1306 CCAAGGAGGACCATGACACTTCATGAATATATATGAAGAAACTA 1355
589 GlyAsnSerTrpGluAlaTrpGluCysLysThrAsnPheArgLysGlnPr 605
1356 GGAAT.....GCTGGCACTCCAAAAA...TTCACAAAGTATCC 1391
605 oGlyArgPheTrpProLeuAspValAspTyrGlyVal...LysLysAlaP 621
1392 CAAAAGTTCTACCCCTGGAGATTGACTACGGCCAGGACAAAGAGCGG 1441
621 roLysArgLysAspIleSerGluMetLysSerSerLeuAlaProGln... 636
1442 TGAAGAACCGACAGTAGTAACTCTGGCCACCAAGTCCATGCTCCCAAGCCA 1491
637 LeuLeuGluLeuMetLysMetLeuPheAsnValGluThrTyrArgAlaAl 653
1492 GTTCAGGACCTATCAACAATGATCTTTGATGTGAAAGTATGCAAGGC 1541
653 aMetMetGluPheGlu***AsnMetSerGluMetProLeuGluLysLeu 670
1542 CCATGGGGGTGTGAGATCAACCTT...CAGATGCCCTTGGGGAAGCTGA 1588
670 er****AsnIleGlu***GlyPheGluAlaLeuThr***** 686
1589 CAAAAGGCAAACTCCAGGCCGGTACTCCATCTCAGGTC..... 1628
687 LeuPheGluGlyHisArg***SerSerThrGlyLeu***GluLysAla** 703
1629CAGCAGGTGGTCTCCAGGCGCAGCAGCACCTCAGAT 1666
703 *LeuLeu*****PheSerLeuLeuSerLeuPheIleL 720
1667 CCTG.....GATCTCTCAAAATCGCTTTTACATCC 1695
720 euIleLeu.....TyrGlyMetArg.....MetIleSer..... 729
1696 TGATCCCCCAGCACTTGGGATGAAGGATCCTCTGCTCTGGAACAATGCA 1745
730TyrSerLysAlaLysMetLeuGluAlaLeuGlnAspIleG1 743
1746 GACAGTGTGCGAGGCCAAGGTAGAAATGCTGCAACACTGCTGCATTTGA 1795
743 uIleAlaSerLysIleVal.....GlyPheAspSerAspSerGluS 758
1796 GGTAGCCTAGCGTCTGCTAGGCGAGGGTCTCACGATAGCAGGAAGGACT 1845
758 erLeuAspAspLysTyrMetLysLeuHisCysAspIleThrProLeuAla 774
1846 CCATCGATGCTCACTGAGAGCTCAAAACTGACATTAAGGTGGTTGAC 1895
775 HisAspSerGluAspTyrLysLeuIleGluLysThrLeuLeuAsnThrH1 791
1896 AGAGATTCTCAAGAGCTGAGATCATCAGGAAGTATGTTAAGAACACTCA 1945
791 ala.Pro...ThrHisLysAspTrpSerLeuGluLeuGluValPhe 806
1946 TGCAACCAACACACACAGGATGATATGACTTGGAAAGTCAATTGATGCTT 1995
807 SerLeuAspArgAspGlyGluLeuAsnLysTyrSerArgTyrLysAsnAs 823
1996 AAGATAGAGTGTGAAGAGGAGTCCAGCACTACAAAGCCCTTTAAG...CA 2042

823 nLeuHisAsnLysMetLeuLeuTrpHisGlySerArgLeuThrAsnPheV 840
2043 GCTTCATAACTGAAGGTGCTGTGGCATGGTCCAGGACCACCAACTTTG 2092
840 aIcIlyIleLeuSerGlnGlyLeuArgIleAlaProGluAlaProVal 856
2093 CTGGGATCTGCTCCCTGGGTCTTTGGATAGCCCTGCTGCAAGCAGCTGTG 2142
857 ThrGlyTyrMetPheGlyLysGlyLeuTyrPheAlaAspLeuValSerly 873
2143 ATGGGCTACATGTTTGTAAAGTATCTATTTCGCTGATCTTGTCTCCAA 2192
873 sSerAlaGlnTyrCysTyrValAspArgAsnAsnProValGlyLeuMetL 890
2193 GAGTGCAACGACTGCCATACATCTTAGGAAGACCAATAGGTTAATCC 2242
890 euLeuSerGluValAlaLeuGlyAspMetTyrGluLeuLysLysAlaThr 906
2243 TGTGCGGAAGAAGTTGCCCTTGGAAACGTGTGTAACCTGAAGCATCTTCA 2292
907 SerMetAspLysProProArgGlyLysHisSerThrLysGlyLeuGly 923
2293 CATATCAGCAAGTTACCCAGGCGACGACAGTGTCAAGGTTGGCAA 2342
923 sThrValProLeuLeuSerGluPheValLysTrpArgAspValValV 940
2343 AACTACTCTGACCTTTCAGCTAGTATCCCA...CTGGATGGTGTAGAGG 2389
940 aIProCysGlyLysProValProSerSerIleArgSerSerGluLeuMet 956
2390 TTCTCTCTGGGACCGAGGTTTTCATCTGTGTGTAATGACACTGTCTACTGT 2439
957 TyrAsnGluTyrIleValTyrAsnThrSerGlnValLysMetGlnPheLe 973
2440 TATATGAGTACATCTCTATGATATGCTCAGGTAAATCTGAAATATCT 2489
973 uLeuLysValArgPheHisHisLys 981
2490 GCTGAAACTGAAATCAATTTTAAG 2514

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-044-618-4

seq_documentation_block:

: Sequence 4, Application US/08044618
: Patent No. 5449605

: GENERAL INFORMATION:

: APPLICANT: SMULSON, MARK

: TITLE OF INVENTION: METHOD OR DETECTING A PREDISPOSITION TO

: TITLE OF INVENTION: CANCER BY THE USED OF RESTRICTION FRAGMENT LENGTH

: TITLE OF INVENTION: POLYMORPHISM OF THE GENE FOR THE HUMAN POLY (ADP-RIBOSE)

: NUMBER OF SEQUENCES: 7

: CORRESPONDENCE ADDRESS:

: ADDRESSEE: Sterne, Kessler, Goldstein & Fox

: STREET: 1225 Connecticut Suite 300

: CITY: Washington

: STATE: D.C.

: ZIP: 20036

: COMPUTER READABLE FORM:

: MEDIUM TYPE: Floppy disk

: COMPUTER: IBM PC compatible

: OPERATING SYSTEM: PC-DOS/MS-DOS

: SOFTWARE: PatentIn Release #1.0, Version #1.25

: CURRENT APPLICATION DATA:

: APPLICATION NUMBER: US/08/044,618

: FILING DATE: 19930406

: CLASSIFICATION: 435

: PRIOR APPLICATION DATA:

: APPLICATION NUMBER: US/07/257,696

: FILING DATE: 14-OCT-1988

: ATTORNEY/AGENT INFORMATION:

: NAME: FOX, SAMUEL L


```
1413 GCAGGATGCTGGAGCGGGTCCTTCACACGCGCAGGAGGTGAGAGG 1452
286 .....ArgCysAlaAspGlyMetLeuPheGlyAlaLeuGlyPro 298
1463 AGCGGGCCAGAGAGCTGAGCGGTCCAGGCTCGCTGATCACACT 1512
299 Cys...ProVal..... 301
1513 GTGAGCCCATGTTGGGCGAGGCGCTGCAAGTCGTTGAGTATGCAATGAA 1562
302 ....CysAlaAsnGlyMetTyrTyrAsnGlyGlnTyrGlnCysSerG 317
1563 GAAATGTCCTCAATGGCATGTC..... 1584
317 LyAsnValSerGluTrpSerLysCysThrTyrSerAlaThrGluProVal 333
1585 .....TCGAGATCAAGTACGATGGA...GAGCGAGTC 1614
334 ArgValLysLysLysTrpGlnIleProHisGlyThrLysAsnAspTyrLe 350
1615 CAGGTGCATAG.....AATGAGACCACTT 1640
350 uMetLysTrpPheLysSerGlnLysValLysLysProGluArgVal.... 365
1641 CAGTACTTCAGCGCGCAGTCTCAAGCGCTTCCTTCACAAAGGTGGCC 1690
366 .....LeuProMetSerProGluLysSerGlySerLys 377
1691 ACTTTAGGACTACATTCCTCCAGCGCTTTCCTGGGGCCACACATGATC 1740
378 AlaThrGlnArgThrSerLeuLeuSerLysGlyLeuAspLysLeuAr 394
1741 TTGGATTCTGAAGTCTTCTGATTGACACAAAGACAGGCAACCACTGCC 1790
394 gPheSerValValGlyGlnSerLysGluAlaAla 405
1791 CTTGGGACTCTGGGAGTACACAAAGAGCGCC 1824
```

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:us-08-464-402-1

seq_documentation_block:

Sequence 1, Application US/08464402

Patent No. 5858705

GENERAL INFORMATION:

APPLICANT: WEI, ET AL.

TITLE OF INVENTION: Human DNA Ligase III

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,

ADDRESSER: CECCHI, STEWART & OLSTEIN

STREET: 6 BECKER FARM ROAD

CITY: ROSELAND

STATE: NEW JERSEY

COUNTRY: USA

ZIP: 07068

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 INCH DISKETTE

COMPUTER: IBM PS/2

OPERATING SYSTEM: MS-DOS

SOFTWARE: WORD PERFECT 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/464,402

FILING DATE: June 5, 1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/03939

FILING DATE: 31 MAR 95

ATTORNEY/AGENT INFORMATION:

NAME: FERRARO, GREGORY D.

REGISTRATION NUMBER: 36,134

REFERENCE/DOCKET NUMBER: 325800-388

TELECOMMUNICATION INFORMATION:

```
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3417 BASE PAIRS
TYPE: NUCLEIC ACID
STRANDEDNESS: SINGLE
TOPOLOGY: LINEAR
MOLECULE TYPE: CDNA
US-08-464-402-1

alignment_scores:
Quality: 173.50 Length: 530
Ratio: 0.735 Gaps: 24
Percent Similarity: 44.528 Percent Identity: 20.189

alignment_block:
US-09-236-995D-2 x US-08-464-402-1 ..
Align seg 1/1 to: US-08-464-402-1 from: 1 to: 3417
```

```
11 GluTyrAlaLysSerGlyArgAlaSerCysLysSerCysArgSerProI 27
358 GACTATGCCAAGCGTGGCACAGTGGCTGCAAAAAAATGCAAGGAAAGAT 407
27 eAlaLysAspGlnLeuArgLeuGlyLysMetVal.....GlnAlaSerG 42
408 TGTGAAGGCGGTATGCCGAATTTGGCAAGTGGTGGCCCAATCCCTTCTCAG 457
42 LnPheAspGlyPheMetProMetTrpAsnHisAlaArgCysIlePheSer 58
458 AGCTCGGGGTGATATGAAGAGTGGTACCACATTAATGTCATGTTTGAG 507
59 Lys.....LysAsnGlnIleLysSerValAspAs 68
508 AAATAGAGCGGGCGGGCCACCACAAAAAATCGAGGACCTCACAGA 557
68 pValGluGlyIleAspAlaLeuArgTrpAspAspGlnGluLysIleArgA 85
558 GCTGAAGGCTGGGAAGAGTGGGAAGATAATGAGAAGGAACAGATAACCC 607
85 snTyrValGlySerAlaSerAlaGlyThrSerSerThrAlaAlaProPro 101
608 AGCACATTGCACATCTGCTTCTTAAGGACAGCAGGTACA.....CCAAAG 651
102 GluLysCysThrIleGluIle..... 108
652 AAGAAAGCTGTTGTCAGGCTAAGTTGACACCACCTGGCCAGGTGACTTC 701
108 ..... 108
702 TCCAGTGAAAGCGGCTCATTTGTCCACCATGACCAATCCCGGAAATTTT 751
109 .....AlaProSerAlaArg 113
752 CTGGCTTTTACCCCAAGCCCACTCTGGGAAGCCCTCGAGGCCCC 801
114 ThrSerCysArgArg.CysSerGluLysIleThrLysGlySerValArg. 129
802 ACCCCCTAAGAGAAGTCTGTCTTCAAGCAAAATGTGACCCCGCAGCATAGGA 851
130 LeuSerAlaLys..... 133
852 CTGCTGTACGGGAGTTTCGAAAGTTATGCCCCATGCTGGCCCGGATAATC 901
134 .....LeuGluSerGluGlyPro.....LysGlyIlePro..... 143
902 CTAGCTACAACACAGGACCCAGATGACCCAGGACTTCCTTCGGAAAGGC 951
144 .....Trp.TyrHisAlaAsnCysPhe..... 150
952 TCAGCAGGAGATGTTTCCACGGTGATGTGTACCTAACAGTGAAGCTGCT 1001
```

```

151 .....PheGluValSerProSerAlaThrV 159
1002 GCTGCCAGGAGCTATTAGAGCTGTTTACAACCTTGAACGATAAGCAGATTG 1051
159 aGluLysPheSerGlyTrpAspThrLeuSerAspGluAspLysArgThr 175
1052 TCAAGCTTTTCAGTCGCATTTTAACTGCAACCCAGCATGATGATGACGCG 1101
176 MetLeuAspLeuValLysLysAspValGlyAsnAsn..... 187
1102 .....GACCTAGACAGCGGTGACGTGTCAGACACAATCAGAGCTTCCTT 1145
188 .GluGlnAsnLys.....GlySerLysArgLysLysSerGluA 200
1146 TCAGCAGAGCAAGTCTTTCCGCCAGCTGCGCAAGAGCTCCTTACCATCC 1195
200 snAspIleAspSerTyrLysSerAlaArgLeuAspGluSerThrSerGlu 216
1196 AGGAAGTGGATGAGTTC...CTTCTGGCGCTGTCCAAGCTCACCAAGGAG 1242
217 GlyThrValArgAsnLysGlyGlnLeuValAspProArgGlySerAsnTh 233
1243 GAT.....GAGCAGACAGCCCTCAGACACATTCGCTCCAGGTG 1283
233 rSerSerAlaAspIleGlnLeuLysLysLeuLysGluGlnSerAspThrLeu 250
1284 TACACCAATGACCTTAATCATCATCAGG..... 1314
250 rPlyLeuLysAspGlyLeuLysThrHisValSerAlaGluLeuArg 266
1315 ..TTGATCAACATCATCTGAAGATGAATCAGTGTGCAAAACATGTGTTA 1362
267 AspMetLeuGluAlaAsnGlyGlnAspThrSerGlyProGluArgHisLe 283
1363 GAGCCCTTGACCCCAATGCTTATGAAGCCTTCAAGCCTCGCGCAACCT 1412
283 uLeuAsp..... 285
1413 GCAGGATGTGGTGAGGGTCTTCAACACGCGCAGGAGTGGAGAAGG 1462
286 .....ArgCysAlaAspGlyMetLeuPheGlyAlaLeuGlyPro 298
1463 AGCCGGCGCCAGAGCAGCTCTGACGCTCCAGGCTCGCTGATGACACT 1512
299 Cys...ProVal..... 301
1513 GTGACGCCCTGTTGGCGGAGCCCTGCAAGTCCGTTGAGTATGCAATGAA 1562
302 ....CysAlaAsnGlyMetTyrTyrTyrAsnGlyGlnTyrGlnCysSerG 317
1563 GAAATGTCCCAATGCGATGTC..... 1584
317 lyAsnValSerGluTrpSerLysCysThrTyrSerAlaThrGluProVal 333
1585 .....TCTGATCAACAGTAGCATGGA...GAGCGAGTC 1614
334 ArgValLysLysLysTrpGlnIleProHisGlyThrLysAsnAspTyrLe 350
1615 CAGGTGCATAG.....ATGGAGACCACTT 1640
350 uMetLysTrpPheLysSerGlnLysValLysLysProGluArgVal.... 365
1641 CAGCTACTTCAGCGCGCAGTCTCAAGCGCTCCTTCTCACAAGGTGGCC 1690
366 .....LeuProMetSerProGluLysSerGlySerLys 377
1691 ACTTTAAGGACTACATTCCTCCAGGCTTTTCTGGGGGCGCACAGCATGATC 1740
378 AlaThrGlnArgThrSerLeuLeuSerSerLysGlyLeuAspLysLeu 394
1741 TTGGATTCTGAAGTCTTCTGATTGACACACAGACAGGCAACCACTGCC 1790
394 gPheSerValValGlyGlnSerLysGluAlaA 405
1791 CTTTGGGACTCTGGGAGTCAACCCGAAAGCAGCC 1824
seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-196-387-1
seq_documentation_block:
; Sequence 1, Application US/09196387
; Patent No. 6277613
; GENERAL INFORMATION:
; APPLICANT: de Lange, Titia
; APPLICANT: Smith, Susan
; TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS
; TITLE OF INVENTION: OF USE THEREOF
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue, 4th Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/196.387
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/095.225
; FILING DATE: June 10, 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-230 CIP1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4134 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; US-09-196-387-1
alignment_scores:
Quality: 157.50 Length: 258
Ratio: 1.382 Gaps: 11
Percent Similarity: 44.186 Percent Identity: 22.868
alignment_block:
US-09-236-995D-2 x US-09-196-387-1
Align seg 1/1 to: US-09-196-387-1 from: 1 to: 4134
763 TyrMetLysLeuHisCys.....AspIleThrProLeuAl 774
3291 TATTTGACTTTTCACTGCTGTAATCAGGAACGATTTTGTGATCTTGC 3340
774 aliAspSerGluAspTyrLysLeuIleGluGlnTyrLeuLeuAsnThrH 791
3341 TCCAGAAGATAAAGATATCAGTCAGTGGGAAGAGAGATGCAAGTACT. 3389
791 isAlaProThrHisLysAspTrpSerLeuGluLeuGluValPheSer 807
|||||
```

```
3390 ..ATTCAGACACAGATGGTGGT..... 3413
808 LeuAspArgAspGlyGluLeuAsnLysTyr..... 817
3414 ...AATGCTGGCGGCATCTTCAACAGATACATGTCATTCGAATTCACAA 3460
817 ..... 817
3461 AGTTGTCAACAAGAAGTTGAGGAGCGGTTCTGCCACCGACAGAAGGAAG 3510
818 ..SerArgTyrLysAsnAsnLeuHisAsnLysMetLeuLeuTrpHisGly 833
3511 TGCTGAGGAGATACACACCATCAATGACCGCATGTGTTCATCGT 3560
834 SerArgLeuThrAsnPheValGlyLeuLeuSerGlnGlyLeuArgIleAl 850
3561 TCTCCTTTTCATTAAT.....GCCATTATTCATAAAGGTTT..... 3596
850 aProProGluAlaProValThrGlyTyrMetPheGlyLysGlyLeuTyrP 867
3597 .GATGAGCGCATGTCATACATAGGAGGAATGTTGGCGCGGATTTATT 3645
867 heAlaAspLeuValSerLysSerAlaGlnTyr..... 877
3646 TTGCTGAAACCTCTCAAAACACACCAATATGTTATGGAAATGGAGGA 3695
878 .....CysTyrValAspArgAs 883
3696 GGAACAGCGTGCCCTACACACAGGAGGTCATGCTATATATGTCACAG 3745
883 nAsnProValGlyLeuMetLeuLeuSerGluValAlaLeuGlyAspMetT 900
3746 ACAG.....ATGCTCTCTGTGTAGAGTGACCCCTGGGAAATCCT 3783
900 yrGluLeuLysLysAlaThrSerMetAspLysProProArgGlyLysHis 916
3784 TTCTGCAGTTAGCACCATGAAATGCGCCCGCCCTCCAGGCCACCCAC 3833
917 SerThrLysGlyLeuGlyLysThrValProLeuGluSerGluPheVal 933
3834 TCAGTC..... 3839
933 strArgAspValValProValProCysGlyLysProValProSerSerI 950
3840 .....ATTGGTAGACCG.....AGCG 3855
950 leArgSerSerGluLeuMetTyrAsnGluTyrIleValTyrAsnThrSer 966
3856 TCAATGGG.....CTGGCATATCTCAATATGTCATCTACAGAGGAA 3899
967 GlnValLysMetGlnPheLeuLeu 974
3900 CAGGCATACCCAGAGATCTATC 3923
```

seq_name: /cgn2_6/plodata/2/ina/6B_COMB.seq:US-09-196-387-7

seq_documentation_block:

; Sequence 7, Application US/09196387

; Patent No. 6277613

; GENERAL INFORMATION:

; APPLICANT: de Lange, Titia

; APPLICANT: Smith, Susan

; TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS

; NUMBER OF SEQUENCES: 12

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Klauber & Jackson

; STREET: 411 Hackensack Avenue, 4th Floor

; CITY: Hackensack

; STATE: New Jersey

; COUNTRY: USA

; ZIP: 07601

; COMPUTER READABLE FORM:

```
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/196.387
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/095,225
FILING DATE: June 10, 1998
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-230 CIP1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 4491 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 6..2027
US-09-196-387-7

alignment_scores:
Quality: 157.50 Length: 258
Ratio: 1.382 Gaps: 11
Percent Similarity: 44.186 Percent Identity: 22.868

alignment_block:
US-09-236-995D-2 x US-09-196-387-7 ..

Align seg 1/1 to: US-09-196-387-7 from: 1 to: 4491

763 TyrMetLysLysLeuHisCys.....AspIleThrProLeuAl 774
3648 TATTGACTTTTCACTGTGTAATCAGGAGACGATTTTGTCTGATCTTGC 3697
774 aHisAspSerGluAspTyrLysLeuIleGluGlnTyrLeuLeuAsnThrH 791
3698 TCCAGAAGATAAAGATATCATGTCAGTGGGAAGAAGAGATGCAAAAGTACT. 3746
791 isAlaProThrHisLysAspTrpSerLeuGluLeuGluValPheSer 807
3747 ..ATTCGAGAACACAGAGATGGTGGT..... 3770
808 LeuAspArgAspGlyGluLeuAsnLysTyr..... 817
3771 ...AATGCTGGCGGCATCTTCAACAGATACATGTCATTCGAATTCACAA 817
817 ..... 817
3818 AGTTGTCAACAAGAAGTTGAGGAGCGGTTCTGCCACCGACAGAAGGAAG 3867
818 ..SerArgTyrLysAsnAsnLeuHisAsnLysMetLeuLeuTrpHisGly 833
3868 TGCTGAGGAGATACACACCATCAATGACCGCATGTGTTCATCGT 3917
834 SerArgLeuThrAsnPheValGlyIleLeuSerGlnGlyLeuArgIleAl 850
3918 TCTCCTTTTCATTAAT.....GCCATTATTCATAAAGGTTT..... 3953
850 aProProGluAlaProValThrGlyTyrMetPheGlyLysGlyLeuTyrP 867
```

```

3954 .GATGACGACATGCATACATAGAGGAAATGTTTGGGGCCGGGATTTATT 4002
867 heAlaAspLeuValSerLysSerAlaGlnTyr..... 877
|||||: ||||||| |||||
4003 TTGCTGAAAACCTCCTCAAAAGCAACCAATATGTTTATGGAATTGGAGGA 4052
878 .....CystyrValAspArgAs 883
|||||: |||||
4053 GGAACAGCTGCCCTACACACAGGACAGGTCATGCTATATATGTCACAG 4102
883 nAsnProValGlyLeuMetLeuSerGluValAlaLeuGlyAspMetT 900
: : |||||: |||||: |||||
4103 ACA.....ATGCTCTTCTGTAGATGACCTTTGGGAAATPCTT 4140
900 yrGluLeuLysLysAlaThrSerMetAspLysProProArgGlyLysHis 916
: : |||||: ||||| ||||| |||||
4141 TTCTGCAGTTTAGCACCATTGAAATGGCCACGCCCTCCAGGGCACACC 4190
917 SerThrLysGlyLeuGlyLysThrValProLeuGluSerGluPheVally 933
|||||: |||||
4191 TCAGTC..... 4196
933 strPArgAspValValValProCysGlyLysProValProSerSerI 950
|||||: |||||
4197 .....ATTGGTAGACCG.....AGCG 4212
950 leArgSerSerGluLeuMetTyrAsnGluTyrIleValTyrAsnThrSer 966
: : |||||: |||||: |||||
4213 TCAATGGG.....CTGGCATATGCTGAATATGTCATCTACAGAGGAGAA 4256
967 GlnValLysMetGlnPheLeuLeu 974
|||||: |||||: |||||
4257 CAGGCATACCCAGACTATCTTATC 4280
seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-196-387-9
seq_documentation_block:
: Sequence 9, Application US/09196387
: Patent No. 6277613
: GENERAL INFORMATION:
: APPLICANT: de Lange, Titia
: APPLICANT: Smith, Susan
: TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS
: TITLE OF INVENTION: OF USE THEREOF
: NUMBER OF SEQUENCES: 12
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Klauber & Jackson
: STREET: 411 Hackensack Avenue, 4th Floor
: CITY: Hackensack
: STATE: New Jersey
: COUNTRY: USA
: ZIP: 07601
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/196,387
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 09/095,225
: FILING DATE: June 10, 1998
: ATTORNEY/AGENT INFORMATION:
: NAME: Jackson Esq., David A.
: REGISTRATION NUMBER: 26,742
: REFERENCE/DOCKET NUMBER: 600-1-230 CIP1
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 201-487-5800
: TELEFAX: 201-343-1684
: TELEX: 133521

```


304 nGlyMetTyrTyrAsnGlyGlnTyrGlnCysSerGlyAsnValSerG 321
1536 1536
321 luTrpSerLysCysThrTyrSerAlaThrGluProValArgValLysLys 337
1537 1537
338 LysTrpGlnIleProHisGlyThrLysAsnAspTyrLeuMetLysTrpPh 354
1582 TTTCATGCAAAATC...CATGCGACCTTCTATGATTCCTCCTAAAGAAG...TT 1625
354 eLysSerGln 357
1626 CACATCCCAA 1635

seq_name: /cgn2_6/ptodata/2/ina/5b_COMB.seq:us-08-685-576-2

seq_documentation block:
: Sequence 2, Application US/08685576
: Patent No. 5906819
: GENERAL INFORMATION:
: APPLICANT: Kaibuchi, Koza
: APPLICANT: Iwamatsu, Akihiro
: APPLICANT: Nakano, Takeshi
: APPLICANT: Ito, Masaaki
: APPLICANT: Takahashi, No. 5906819uaki
: TITLE OF INVENTION: RHO TARGET PROTEIN RHO-KINASE
: NUMBER OF SEQUENCES: 16
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Foley & Lardner
: STREET: 3000 K Street, N.W., Suite 500
: CITY: Washington
: STATE: D.C.
: COUNTRY: USA
: ZIP: 20007-5109
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/685,576
: FILING DATE: 24-JUL-1996
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 7-325129
: FILING DATE: 20-NOV-1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 8-17150
: FILING DATE: 05-JAN-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 8-131206
: FILING DATE: 26-APR-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Bent, Stephen A.
: REGISTRATION NUMBER: 29,768
: REFERENCE/DOCKET NUMBER: 16887/843
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202)672-5300
: TELEFAX: (202)672-5399
: TELEX: 904136
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 5053 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: ORIGINAL SOURCE:
: ORGANISM: Bovine
: FEATURE:
: NAME/KEY: CDS

LOCATION: 1..4164
US-08-685-576-2
alignment_scores:
Quality: 113.50 Length: 949
Ratio: 0.268 Gaps: 42
Percent Similarity: 44.573 Percent Identity: 18.230
alignment_block:
US-09-236-995D-2 x US-08-685-576-2 ..
Align seg 1/1 to: US-08-685-576-2 from: 1 to: 5053
20 CysLysSerCysArgSerProIleAlaLys.....As 30
1408 TGAAGTCTCTTAACTCGCTTAGAGAAAGTGGCAAGGAGTTAGAA 1457
30 pGlnLeuArgLeuGlyLysMetValGlnAlaSer.....GlnPheAspG 45
1458 AGAGATTACCTTAAAGGAAAAATGTGAATCAACATTAGACAAATTAGAAA 1507
45 LysPheMetProMetTrpAsnHisAlaArgCysLlePheSerLysLys... 60
1508 GAGAAAAAGCAGCTTCTTCAGCACAAAAATGCGAGATATCAGCGGAAGCT 1557
61AsnGlnIleLysSerVa 66
1558 GATCATGAACAGACACAAAGGAAATTTGGAGAAATGATGTTAACAGTTT 1607
66 lAspAspValGluGlyIleAspAlaLeuArgTrpAspGlnGluLysI 83
1608 AAAAGAT.....CAGCTTGAAGATTGAAAAA. 1635
83 leArgAsnTyrValGlySerAlaSerAlaGlyThrSerSerThrAlaLa 99
1636 ..AGAAATCAGAACTCTCAGATATCCACT..... 1662
100 ProProGluLysCysThrIleGluIleAlaProSerAlaArgThrSerCy 116
1663GAGAAAGTGAATCAACTCCAGAGACAACTGGATGAAACCAATGC 1706
116 sArgArgCysSerGluLysIleThrLysGlySerValArgLeuSerAlaL 133
1707 TTTGCTCGCAACAGAAATCTGATCTCGAGCCCGGTTAAGG.....A 1747
133 ysLeuGluSerGluGlyProLysGlyIleProTrpTyrHisAla..... 147
1748 AAACACAGGAGAAAGTTCAAACAGATTTCAGCAGCTGGAATCTAACAA 1797
148AsnCysPhePheGluValSerProSerAlaTh 158
1798 AGAGATCTACAAGACAAAAATTCCTGCTGGAGACTGCCCAAGTTAAACT 1847
158 rValGluLysPheSerGlyTrpAspThrLeuSerAspGluAspLysArgT 175
1848 TGAAGAGAAATTTATCAATCTTCAGTCAGTTCTAGATCTGGAAGGAGG. 1896
175 hrMetLeuAspLeuValLysLysAspValGlyAsnGlnGluAsnLys 191
1897GACCGAACCCAT 1908
192 GlySerLysArgLysLysSerGluAsnAspIleAspSerTyrLysSerAl 208
1909 GGATCA.....GAGATTATTAATGATTACAGGT...AGAATATC 1946
208 aArgLeuAspGluSerThrSerGluGlyThrVal..... 219
1947 TGGCTAGACAGACAGATGTAAGAATGTAAATCTTTATTAGCAAAAGTAG 1996
220 ..ArgAsnLysGlyGlnLeuValAsp.....ProArgGly 230
1997 AGCTGGAGAGAGACAACTACAGGAGAGATTTACTGATTGGAAAAAGGAA 2046

181	ysLysaspValGlyAsnAsnGluInAsnLysGlySerLysArgLysLys	197
1182	AAAAA.....TATGAAATGAGCATCAGGTGGTAGTACGAAAACGG	1225
198SerCluAsnAspIleaspSerTyrLysSerAlaAr	209
226	GATGCAGGTGGTACAACTACTACTACTATTTATGATGATGATGAAAAAATTT	1275
209	gLeuAspGluSerThrSerGluGlyThrValArgAsnLysGlyGlnLeu	226
1276	TTATGACGAACCTAAT.....AAAAGTGAATATAGAACCG	1310
226	aIAspProArgGlySerAsnThrSerSerAlaAspIleGlnLeuLysLeu	242
1311	TTGATAAATTTTGGAAAAATTAAGTAATGAAGAATAATGACAAAAGTT	1360
243	Lys...GluGlnSerAspThrLeuTrpLysLeuLysAspGlyLeuLysTh	258
1361	AAAGACGAAGAGGAGCAATTT.....1384	
258	rHisValSerAlaAlaGluLeuArgAspMetLeuGluAlaAsnGlyGlnA	275
1385GATTTTAAAAACCGTTAATAGTGTATA1409	
275	spThrSerGlyProGluArgHisLeuLeuAspArgCysAlaAspGlyMet	291
1410	GTACTAGTGGT.....GCTAGTGGCACT	1432
292	LeuPheGlyAlaLeuGly.....ProCys	299
1413	AATGTTGAAAGTCAAGAACATTTTATCTGTTCAAAATATGCGCAACCTG	1482
299	sProValCysAlaAsnGlyMetTyrTyrAsnGlyGlnTyrGlnCysS	316
1483	CCCTTATTTGT.....GGAGTGAAGAAGGTAAATAAT.....G	1514
316	erGlyAsnValSerGluTrp.....SerLysCysThrTyr	327
1515	GTGGTAGTACTAATGAATGGAGAGAAAAAATAATGGCAAGTGC.....	1558
328	SerAlaThrGluProValArgValLysLysLysTrpGlnIleProHisGl	344
15581558	
344	yThrLysAsnAspTyrLeuMetLysTrpPheLysSerGlnLysValLysL	361
1559AAGGTGGAACACTTATG	1577
361	ysProGluArgValLeuProMetSerProGluLysSerGlySerLys	377
1578	AGCCTAAA.....CCCGACAAAGAAGT.....1600	
378	AlaThrGlnArgThrSerLeuLeuSerSerLysGlyLeuLysLysLeuAr	394
1601	...ACTACTATTACAATCCCTTAAGAGTGGTAAGGACATGATGATATGA	1647
394	gPheSerVal.....397	
1648	ACAAATAATTAACCAAAATTTTGTGTGATGAAAAAATGTTGATACATAAATA	1697
398valGlyGlnSerLysGluAlaAlaAsn	406
1698	GTGGTGGTAGTGGTACCGGTGGTAGTGGTGGTGAACATGGTGACACAG	1747
407	GluTrpIleGlyLysLeuLysLeuAlaGlyAlaAsnPheTyrAlaArgVa	423
1748	GAATTCGTATGAAGATGCAATGTTATAAAGGTCAACATGTAGTGAAGT	1797
423	IvalLysAspIleAspCysLeuIleAlaCysGlyGluLeuAspAsnGluA	440
1798	TGACACCATCAGGAT.....GACGACGAGGATTATGAAA	1832

```
seq_documentation_block:
  Sequence 13, Application US/08487826B
  Patent No. 5993827
  GENERAL INFORMATION:
    APPLICANT: Sim, Kim L.
    APPLICANT: Chitnls, Chetan
    APPLICANT: Miller, Louis H.
    APPLICANT: Peterson, David S.
    APPLICANT: Su, Xin-zhaun
    APPLICANT: Wellens, Thomas E.
  TITLE OF INVENTION: BINDING DOMAI
```

TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:

CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe Martens Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: California
COUNTRY: US

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,826B
FILING DATE: 10-SEP-1993
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Israelsen, Ned
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NHI21-001CP1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176

INFORMATION FOR SEQ ID NO: 13:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 19124 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 -487-8268-13

```

alignment_scores:
  Quality: 112.00      Length: 576
  Ratio: 0.475        Gaps: 28
  Percent Similarity: 40.972  Percent Identity: 19.271

alignment_block:
  US-09-236-995D-2 x US-08-487-826B-13  ..

  Align seg 1/1 to: US-08-487-826B-13 from: 1 to: 19124

148 AsnCysphePheGluValSerProSerAlaThrValGluLysPheSerG1 164
   :::::||||:::      ::|||  |||||:::
8408 AGCTGTTTGATGTCATGTAUUCT.....TAGCTTGATGGTAAATAA 8451

164 yTrpAspThrLeuSerAspGluAspLysArgThrMetLeuAspLeuVal 181
   ||:::||||:::      ||:::||||:::
8452 CCAAAAGAACAAATTTGCAACACAGAAAAAAAATATGATGAAGAAATAA 8501

181 yLysAspValGlyAsnAsnGluGlnAsnLysGlySerLysArgLysLys 197
   |||||.....      ::|||  |||||:::
8502 AAAAA.....TATGAAATGAGCATCAGGTGGTAGTAGGCAAAACCG 8545

198 .....SerGluAsnAspIleAspSerTyLysSerAlaAr 209
   :::  :::::  |||||:::
8546 GATCGAGTGTGACAACTACTACTAATTATGATGGATATGAAAAAAATT 8595

209 gLeuAspGluSerThrSerGluGlyThrValArgAsnLysGlyGlnLeuV 226
   |||||  :::::  |||||:::
8596 TTAGCGACGACCTAAT.....AAAAGTGAATATAGAACCG 8630

226 aAspProArgGlySerAsnThrSerSerAlaAspIleGlnLeuLysLeu 242
   |||||  :::::  |||||:::
8631 TTGATAAATTTTCGAAAAAATTAGTAATGACAAATATGCACAAAGTT 8680

243 Lys...GluGlnSerAspThrLeuTrpLysLeuLysAspGlyLeuLysTh 258
   |||  ||:::  ||||:

```

```

520 s***LeuLysHis.....ValLeuThr***His***ValCysThrG 534
    |||:||||:|
9331 ATGTTTACAAATGTAACAGATAAATGTGGAAACAAATAATGT.... 9376
    |||:||||:|
534 LyTyrTyrValLeuGlnIleIleGluGlnAspGlySerGluCysTyr 550
    |||:||||:|
9377 .....ATAATGATTGTGAA 9391
    |||:||||:|
551 ValPheArgLysTrpGlyArgValGlySerGluLysIleGlyGlnLy 567
    |||:||||:|
9392 TCTTTTAAAGATGATTACA.....CAAAA 9417
    |||:||||:|
567 sleuGluGluMetSerLysThrGluAlaIleLysGluPheLysArgLeuP 584
    |||:||||:|
9418 AAAAGACGAATGGGGAAA.....ATAGTACACATTTTAAACGCCAA 9461
    |||:||||:|
584 heLeuGluLysThrGlyAsnSerTrpGluAlaTrpGluCysLysThrAsn 600
    |||:||||:|
9462 ATATTAAAGGTAGAGGAGTACT..... 9484
    |||:||||:|
601 PheArgLysGlnProGlyArgPheTyrProLeuAspValAspTyrGlyVa 617
    |||:||||:|
9485 .....GACAATACGCCAGATAATCCCATTTGATCAGGATTAATGTCT 9528
    |||:||||:|
617 lLysLysAlaProLysArg.....LysAspIles 627
    |||:||||:|
9529 TCATACAAATTTGCAAGAGAGAAATTTTGAAGGCGATTCCGAAGACGCTT 9578
    |||:||||:|
627 erGluMetLysSer.....SerLeuAlaProGlnLeuGluLeuMet 641
    |||:||||:|
9579 CCGAAGAAAAATCCGAAAAATAGTCTGGATGCAGAGGAGGACAGGAACTA 9628
    |||:||||:|
642 LysMetLeuPheAsnValGluThrTyrArgAlaAlaMetMetGluPheGl 658
    |||:||||:|
9629 AAACACCTTCGCGAAATC.....ATTGAAAGTGA 9657
    |||:||||:|
658 u***AsnMetSerGluMetProLeuGly 667
    |||:||||:|
9658 AGACAATAATCAAGAAGCATCTGTTGGT 9685
    |||:||||:|

```

OM of: US-09-236-995D-2 to: EST:* out_format : pfs

Date: Mar 9, 2002 8:22 PM

About: Results were produced by the GenCore software, version 4.5.
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODE=frame+p2n_model -DEV=xlp
-O=/cgn2_1/USPTO_spool/US09236995/runat_08032002_113148_25093/app_query.fasta_1.1273
-DB=EST -QFMT=fastap -SUFFIX=rst -CAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPEL=0.000 -LOOPEXT=0.000 -GAPEXT=4.500
-GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-FGAPEXT=7.000 -START=1 -MATRix=blomsum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NOHM=ext -MINLEN=0
-MAXLEN=2000000000 -USER=US09236995 -CGCN1_1_5291 -NCPU=6
-ICPU=3 -LONGLOG -NO_XLPY -WAIT -THREADS=1

Search information block:

Query: US-09-236-995D-2
Query length: 982
Database: EST:*
Database sequences: 11351937
Database length: 1077921985
Search time (sec): 2336.950000

score_list:

Sequence	Strd	Orig	ZScore	EScore	Len	Documentation
gb_est1:AW066305	+	856.00	1405.08	4.2e-69	653	! AW066305 687009D02.y1 687 - Ear
gb_est1:AV833309	+	783.00	1283.03	2.7e-62	714	! AV833309 AV833309 K. Sato unpub
gb_est2:BF255013	+	714.00	1170.17	5.1e-56	607	! BF255013 HVSMBF0005L18f Hordeu
gb_est2:BG280821	+	676.00	1103.58	2.6e-52	861	! BG280821 602401166F1 NIH_MGC_20
gb_est1:BEA19432	+	660.00	1081.81	4.3e-51	536	! BEA19432 WMS012.AL2R000101 ITEC
gb_est1:AA401836	+	658.00	1078.25	8.8e-51	669	! AA401836 zV66b12.r1 Soares_tola
gb_est2:BI093436	+	650.00	1059.81	7.2e-50	916	! BI093436 602859629F1 NIH_MGC_10
gb_est1:AV834168	+	642.00	1049.75	2.6e-49	661	! AV834168 AV834168 K. Sato unpub
gb_est2:BI091452	+	642.00	1045.82	4.4e-49	990	! BI091452 602299739F1 NIH_MGC_87
gb_est2:BG429402	+	618.00	1007.65	5.8e-47	834	! BG429402 602858931F1 NIH_MGC_10
gb_est2:BE941860	+	589.00	993.85	3.1e-46	632	! BE941860 602499624F1 NIH_MGC_75
gb_est2:BG751755	+	586.00	966.47	1.1e-44	419	! BG751755 602730481F1 NIH_MGC_43
gb_est2:BF976506	+	581.00	943.33	2.8e-44	771	! BF976506 602244411F1 NIH_MGC_4
gb_est1:AA212857	+	577.50	943.48	2.2e-43	617	! AA212857 mw84d01.r1 Soares_mou
gb_est2:BF042005	+	575.00	940.31	3.3e-43	560	! BF042005 BP250013A20B10 Soares
gb_est1:AI455188	+	573.50	935.17	6.3e-43	728	! AI455188 LD21673.3prime LD Dros
gb_est2:BF101770	+	569.00	934.56	6.8e-43	369	! BF101770 WMS04.C11R000101 ITEC
gb_est2:BG533818	+	567.00	925.10	2.3e-42	678	! BG533818 602562491F1 NIH_MGC_20
gb_est1:BE729084	+	554.00	901.96	4.5e-41	792	! BE729084 602562491F1 NIH_MGC_76
gb_est2:BF459972	+	552.00	900.06	5.7e-41	688	! BF459972 071B10 Mature tuber la
gb_est1:BE516130	+	542.50	884.51	4.2e-40	674	! BE516130 WHE612.D07_H14ZA Whea
gb_est1:BE488132	+	541.00	886.62	3.2e-40	427	! BE488132 177421 BARC 5BOV Bos t
gb_est2:BG281447	+	540.50	884.22	4.4e-40	499	! BG281447 602401961F1 NIH_MGC_20
gb_est2:BF026745	+	537.00	871.65	2.2e-39	975	! BF026745 601671953F1 NIH_MGC_20
gb_est1:AA397988	+	536.00	875.25	1.4e-39	579	! AA397988 t86c12.r1 Soares_tes
gb_est2:BF793705	+	535.50	869.40	2.9e-39	952	! BF793705 602255121F1 NIH_MGC_84
gb_est2:BI217470	+	530.50	863.98	5.8e-39	716	! BI217470 602333675F1 NIH_MGC_1
gb_est2:BG259918	+	530.00	860.58	9.0e-39	924	! BG259918 602371878F1 NIH_MGC_93
gb_est1:BE304453	+	529.50	864.61	5.4e-39	571	! BE304453 601105502F1 NIH_MGC_15
gb_est2:BF702243	+	526.50	859.35	1.1e-38	587	! BF702243 601662539F1 NIH_MGC_15
gb_est1:BE382739	+	526.50	858.57	1.2e-38	634	! BE382739 AV702243 ADB Homo sap
gb_est2:BG16404	+	522.50	850.81	3.2e-38	709	! BG16404 601297575F1 NIH_MGC_18
gb_est2:BG16404	+	521.50	850.71	3.2e-38	607	! BG16404 dad35f02.y1 Wellcome
gb_est2:BI079902	+	521.50	847.39	4.9e-38	844	! BI079902 602867112F1 NIH_MGC_60
gb_est2:BI253230	+	509.50	829.01	5.2e-37	725	! BI253230 602876096F1 NIH_MGC_1
gb_est2:BG287484	+	506.50	822.53	1.2e-36	841	! BG287484 602382082F1 NIH_MGC_12
gb_est2:BG675913	+	503.50	816.56	2.6e-36	928	! BG675913 602622263F1 NIH_MGC_9
gb_est1:AA450493	+	501.50	818.93	1.9e-36	528	! AA450493 vff7907.r1 Soares_mamm

gb_est2:BG698037 + 498.00 808.31 7.4e-36 851 ! BG698037 602659183F1 NCI_CGA
gb_est1:AW403555 + 495.50 810.95 5.2e-36 434 ! AW403555 UI-BF-BK0-abf-e-12-
gb_est2:BG042229 + 493.00 805.33 1.1e-35 502 ! BG042229 su93c12.y1 Gm-c1055
gb_est2:BG476961 + 492.00 800.51 2.0e-35 687 ! BG476961 602522176F1 NIH_MGC

seq_name: gb_est1:AW066305

seq_documentation_block: 653 bp mRNA 30-MAR-2000
LOCUS AW066305
DEFINITION 687009D02.y1 687 - Early embryo from Delaware Zea mays cDNA, mRNA
sequence.
ACCESSION AW066305
VERSION AW066305.1 GI:6021377
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 653)
AUTHORS Walbot V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
UNIVERSITY
JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 687009 row: D column: 02.
FEATURES
Location/Qualifiers
1..653
/organism="Zea mays"
/cultivar="Illinois High Oil"
/db_xref="taxon:4577"
/clone_lib="687 - Early embryo from Delaware"
/tissue_type="embryo"
/dev_stage="14, 21, 28, and 35 days after pollination"
/lab_host="E. coli SOLR"
/note="Organ: embryo; Vector: pBluescript SK; Site: 1: XhoI
; Site 2: EcoRI; Library was prepared by Statagene using
the Uni-ZAP XR system (Stratagene BN937328-12). Clones
were picked by a Q-bot after blue/white selection
(ampicillin resistance - use 100 micrograms/microliter).
Developed from a pool of equal amounts of RNA from
pollinating embryos sampled at 14, 21, 28 and 35 days after
pollination of the Illinois High Oil Maize Strain Cycle
90. This closed strain has been selected for high oil
concentration for 90 generations and originates from the
1890s era open pollinated variety Burr's White"

BASE COUNT 215 a 98 c 161 g 178 t
ORIGIN

alignment_scores:
Quality: 856.00 Length: 213
Ratio: 4.412 Gaps: 0
Percent Similarity: 91.080 Percent Identity: 90.141

alignment_block:
US-09-236-995D-2 x AW066305
Align seg 1/1 to: AW066305 from: 1 to: 653
535 TyrTyrValLeuGlnIleIleGluAnAspGlySerGluCysTyrVa 551
18 TACTATGTACTCCAGATCAATGACAGATGATGGTCTGAGTGCTACGT 67
551 lPheArgLysTrpGlyArgValGlySerGluLysIleGlyGlyGlnLysL 568
68 ATTTCGTAAGTGGGACGGGTGGGAGTGAATAATGGAGGCAAAAC 117

568 euGlucMetSerLysThrGluAlaLeuLysGluPheLysArgLeuPhe 584
|||||
118 TGGAGGAGATGTCACAAACTGAGGCAATCAAGGAATTCACAAAGATATTT 167
585 LeuGluLysThrGlyAsnSerTrpGluAlaTrpGluCysLysThrAsnPh 601
|||||
168 CTTGAGAGAGACTGGAACTCACTGGGAAGCTGGGAATGTAACCAATTT 217
601 eaGlyLysGlnProGlyArgPheTyPrProLeuAspValAspTyPrGlyValL 618
|||||
218 TCGGAAGACCTGGGAGATTTTACCACCTTGATGTTGATGATGTTGTTA 267
618 yslYsAlaProLysArgLysAspLeSerGluMetLysSerSerLeuAla 634
|||||
268 AGAAGACACCAAGGAAAGATATCACTGAAATGAAAGTTCTCTGCT 317
635 ProGlnLeuLeuGluMetLysMetLeuPheAsnValGluThrTyAr 651
|||||
318 CCTCAATGCTAGACTCATGAGATGCTTTTCAATGTTGGAGACATATAG 367
651 gAlaAlaMetMetGluPheGlu***AsnMetSerGluMetProLeuGlyL 668
|||||
368 AGCTGCTATGATGGAATTTGAAATTAATATGTCAGAAATGCCCTCTGGGA 417
668 yslLeuSer*****AsnLeGlu***GlyPheGluAlaLeuThr***** 684
|||||
418 AGCTAAGCAAGGAAATTTGAGAAGGATTTGAAGCATTAACACTGAGATA 467
685 *****LeuPheGluGlyHisArg***SerSerThrGlyLeu***GluLy 701
|||||
468 CAGAAATTTATT.GAAGGACACCGCTCATCAAGCACTGGCTG.TTAGAGAA 515
701 sAla***LeuLeu*****PheSerLeuLeuSerLeuLeuP 718
|||||
516 AGCTTAATGTTGCTCGAGCAATCGCTTTTCACTCTTATCCCTCTAT 565
718 heileLeuLeuLeuTyGlyMetArgMetIleSerTyPrSerLysAlaLys 734
:::
566 TATN.TCTCATATATACGGATGAGGATGATTTG.ATGATCAAGCGGAA 613
735 MetLeuGluAlaLeuGlnAspIleGluIleAlaSerLys 747
|||||
614 ATGCTTTGAAGCTCTGCAGGATATTTGAATTTGCTTCAAAG 652
seq_name: gb_est1:AV833309
seq_documentation_block:
LOCUS AV833309 714 bp mRNA EST
DEFINITION AV833309 K. Sato unpublished cDNA library: Hordeum vulgare subsp.
vulgare shoots germination Hordeum vulgare subsp. vulgare cDNA
clone bags5K04, mRNA sequence.
ACCESSION AV833309
VERSION AV833309.1 GI:14525398
KEYWORDS EST.
SOURCE Hordeum vulgare subsp. vulgare.
ORGANISM Hordeum vulgare subsp. vulgare.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
; Triticeae; Hordeum.
REFERENCE 1 (bases 1 to 714)
Sato,K.
Barley EST sequencing project in NIG and Okayama Univ
Unpublished (2001)
Contact: Kazuhiro Sato
Research Institute for Bioresources
Okayama University, Barley Germplasm Center
Chuo 2-20-1, Kurashiki, Okayama 710-0046, Japan
Email: kzsato@rib.okayama-u.ac.jp,
URL:http://www.rib.okayama-u.ac.jp/barley/
Sato,K., Saisho,D., Takeda,K., Shini,T. and Kohara,Y. Direct
submission:
database:http://www.shigen.nig.ac.jp/barley/Barley.html.

FEATURES
Source
Location/Qualifiers
1..714
/organism="Hordeum vulgare subsp. vulgare"
/cultivar="Haruna NiJo"
/db_xref="taxon:112509"
/clone="bags5K04"
/clone_lib="K. Sato unpublished cDNA library: Hordeum
vulgare subsp. vulgare shoots germination"
/tissue_type="shoots"
/dev_stage="germination"
BASE COUNT 249 a 121 c 162 g 182 t
ORIGIN
alignment_scores:
Quality: 783.00 Length: 240
Ratio: 4.015 Gaps: 4
Percent similarity: 81.250 Percent identity: 71.667
alignment_block:
US-09-236-995D-2 x AV833309 ..
Align seg 1/1 to: AV833309 from: 1 to: 714
509 HisIlePro*****TrpGluLysHisIleGlnCys***LeuLysHisVa 525
|||||
6 CACGAGGCTGGAAGATGCGCAAAAGCATTTACAAATACCAACCCCTAAACAT.. 53
525 lLeuThr***His***ValCys.ThrGlyTyPrTyPrValLeuGlnIleile 541
:::
54 .TTCGTGACATGACACAAAGGTGTTAACAGCTACTATATATCTTCAGATCATC 102
|||||
542 GluGlnAspAspGlySerGluCysTyPrValPheArgLysTrpGlyArgVa 558
|||||
103 GAAGAGGATGATGGAGTGAATGCTATGTTTCGAAAGTGGGGCGAGT 152
558 lGlySerGluLysIleGlyGlnLysLeuGluGluMetSerLysThrG 575
|||||
153 TGGCAGTGAAGGATTTGGTGGAAAGAACTGGAGGAGATGTCAAAAACTG 202
575 luAlaIleLysGluPheLysArgLeuPheLeuGluLysThrGlyAsnSer 591
:::
203 ACCTAATACATGAATTTTAAAGATTTTCGGAAGAGACTGGAAACCCC 252
592 TrpGluAlaTrpGluCysLysThrAsnPheArgLysGlnProGlyArgPh 608
|||||
253 TGGGAAGCATGGGAACAAACAAATTTTCAGAAAGCAGCCTGGGAGATT 302
608 eTyPrProLeuAspValAspTyPrGlyValLysLysAlaProLysArgLysA 625
|||||
303 TTATCCACTTGACATTTGATTACGGAGTTAAGCAAGCACCAGCAAGGAAAG 352
625 spIleSerGluMetLysSerSerLeuAlaProGlnLeuLeuGluLeuMet 641
|||||
353 ACATCAGCAAAATGAAAGATTCACCTGCTCCTCAGGTGCTGGAACCTCATG 402
642 LysMetLeuPheAsnValGluThrTyPrArgAlaAlaMetMetGluPheGl 658
|||||
403 ATGATGCTTTCAATGTTGAAACATATATAGGCTGCTATGATGGAATTTGA 452
658 u***AsnMetSerGluMetProLeuGlyLysLeuSer*****AsnIleG 675
|||||
453 AATCAATATGGCAAAATGCCCTTTGGGAATTTAAGCAAGGAAATATACC 502
675 lu***GlyPheGluAlaLeuThr*****LeuPheGluGlyHis 691
:::
503 AGAAGAGGATTTGAACCATTAATGAGATACAAATCTACTGAT...GAC 549
692 Arg***Ser.SerThrGlyLeu***GluLysAla***LeuLeu***** 708
:::
550 ACTGCAATCAAGAACTGGCTCTTAGAGAGAGCTTGATTTGTTGTCGAAG 599
708 *****PheSerLeuLeuSerLeuPheIleLeuLeuTyPrGly 724


```

882 rgAsnAsnProValGlyLeuMetLeuLeuSerGluValAlaLeuGlyAsp 898
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
111 GAAAAAATCCGACGGCTTGATGCTTCTTCTGAGGTGCTCTAGGAGAC 160
899 MetTyrGluLeuLeuLysAlaThrSerMetAspLysProProArgGlyLys 915
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
161 ATGATGAAGTGAAGAGGCGGACGCAATGGACAACTCCCAAGAGGAAA 210
915 sHisSerThrLysGlyLeuGlyLysThrValProLeuGluSerGluPheV 932
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
211 GCATTCGACCAAGGCTTAGGAAAACTGTGCGCTAGAGTCGGAGTTTG 260
932 allYsTrpArgAspValValProCysGlyLysProValProSer 948
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
261 TGAATGGAGGGATGATATCGTGTGCTGCTGTGGCAAGCAGTGCACGCA 310
949 SerIleArgSerSerGluLeuMetTyrAsnGluTyrIleValTyrAsnTh 965
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
311 TCTATCANGCATCTGAGGCTTGTACACAGGATATATAGTCTACAAAC 360
965 rSerGlnValLysMetGlnPheLeuLeuLysValArgPheHisHisLysA 982
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
361 AGCCAGGTGAAGATGATGCTTGTGAGGTGAGTTCAGATTCGCTCACAAGC 410
982 rg 982
||
411 GT 412

```

seq_name: gb_est1:AA401836

seq_documentation_block: 669 bp mRNA EST 16-MAY-1997
 LOCUS AA401836
 DEFINITION Zv66b12:1 Soares_cotal_fetus_Nb2HF8_9w Homo sapiens cDNA clone IMAGE:758591 5' similar to gb:J03473 NAD(+) ADP-RIBOSYLTRANSFERASE (HUMAN); mRNA sequence.

ACCESSION AA401836
 VERSION AA401836.1 GI:2055855
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 669)
 Hillier, L., Allen, M., Bowles, L., Dubouque, T., Geisel, G., Jost, S., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Stepien, M., Tan, F., Theising, B., White, Y., Wyllie, T., Waterston, R. and Willson, R.
 WashU-Merck EST Project 1997
 Unpublished (1997)
 Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu

TITLE
 JOURNAL
 COMMENT
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.lnl.gov) for further information.
 Seq primer: -28m13 rev2 ET from Amersham
 High quality sequence stop: 493.
 Location/Qualifiers
 1. .669
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:758591"
 /clone_lib="Soares_total_fetus_Nb2HF8_9w"
 /dev_stage="8-9 weeks"
 /lab_host="DH10B"
 /note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from pooled 8-9 week (total) fetus material with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGCGGCGCTTAATTTTTTTTTTTT 3'].

FEATURES
 source

Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7m3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo. *

BASE COUNT 191 a 152 c 157 g 169 t
 ORIGIN

alignment_scores:
 Quality: 658.00 Length: 223
 Ratio: 3.500 Gaps: 2
 Percent Similarity: 84.305 Percent Identity: 55.605
 alignment_block:
 US-09-236-995D-2 x AA401836

Align seg 1/1 to: AA401836 from: 1 to: 669

```

759 LeuAspAspLysTyrMetLysLeuHisCysAspIleThrProLeuAlaHi 775
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2 ATCATGTCACACTATGAGAAGCTCAAACTGACATTAAGGTGGTTGACAG 51
775 sAspSerGluAspTyrLysLeuIleGluGlnTyrLeuLeuAsnThrHisA 792
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
52 AGATTCTGAAGAAGCGGAGATCATCAGGAAGTATGTTAAGAACACTCATG 101
792 laProThrHisLysAspTrpSerLeuGluLeuGluValPheSerLeu 808
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
102 CAACCACACACAACTGCTATGATGTTGGAAGTCATCGATATCTTTAAGATA 151
809 AspArgAspGlyGluLeuAsnLysTyrSerArgTyrLysAsnAsnLeuHi 825
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
152 GAGCGTGAAGGCGATGATGCTGTTACAGCCCTTTAAG...CAGCTTCA 198
825 sAsnLysMetLeuLeuThrHisGlySerArgLeuThrAsnPheValGlyI 842
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
199 TAACCGAAGATTGCTGTGGCAGCGGTCCAGGACCACCAACTTTGCTGGGA 248
842 leLeuSerGlnGlyLeuArgIleAlaProProGluAlaProValThrGly 858
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
249 TCTGTCTCCAGGGTCTTCGGATAGCCCGCCCTGAAGCGCCGTCACAGGC 298
859 TyrMetPheGlyLysGlyLeuTyrPheAlaAspLeuValSerLysSerAl 875
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
299 TACATGTTTGTAAAGGATCTATTCGCTGACATGCTCTCCAGAGAGTGC 348
875 aGlnTyrCysTyrValAspArgAsnAsnProValGlyLeuMetLeuLeu 892
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
349 CAACTACTGCCATACGCTCTCAGGAGACCAATAGGCTTAATCTCTGTGG 398
892 erGluValAlaLeuGlyAspMetTyrGluLeuLysLysAlaThrSerMet 908
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
399 GAGAAGTTGCCCTTGGAAACATGATGATGAAGCAGCAGCTTCACATATC 448
909 AspLysProProArgGlyLysHisSerThrLysGlyLeuGlyLysThrVa 925
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
449 AGCAAGTTACCAAGGCAAGCAGAGTGTCAAAGGTTTGGGCAAACTAC 498
925 lProLeuGluSerLysPheValLysTrpArgAspValValProC 942
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
499 CCCTGATCCTTCAGCTAACATAGT...CTGGATGTTGAGAGCTTCCTC 545
942 ysGlyLysProValProSerSerIleArgSerSerGluLeuMetTyrAsn 958
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
546 TTGGGACCGGGATTTTCATCTGGTGTGAATGACACCTCTCTACTATATAAC 595
959 GluTyrIleValTyrAsnThrSerGlnValLysMetGlnPheLeuLeu 975
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
596 GAGTACATTTCTCTATGATATTTGTCAGGTAATCTCTGAAGTATCTGCTGAA 645
975 sValArgPheHisLys 981
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

```
646 ACTGAATTCRAATTTTAAG 664
seq_name: gb_est2:BI093436
seq_documentation_block:
LOCUS      BI093436      916 bp      mRNA      20-JUN-2001
DEFINITION 602859629f1 NIH_MGC_10 Homo sapiens cDNA clone IMAGE:5000911 5',
            mRNA sequence.
ACCESSION  BI093436
VERSION    BI093436.1 GI:14511766
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 916)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: Incyte Genomics, Inc.
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LAM11033 row: i column: 08
            High quality sequence stop: 837.
            Location/Qualifiers
                1..916
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:5000911"
                /clone_lib="NIH_MGC_10"
                /cell_line="MGC36"
                /lab_host="DH10B"
                /note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI;
                Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt;
                Average insert size 1.5 kb. Library prepared by Life
                Technologies."
BASE COUNT 248 a 209 c 230 g 229 t
ORIGIN
alignment_scores:
    Quality: 650.00      Length: 217
    Ratio: 3.514        Gaps: 2
    Percent Similarity: 85.253      Percent Identity: 56.221
alignment_block:
US-09-236-995D-2 x BI093436 ..
Align seg 1/1 to: BI093436 from: 1 to: 916
765 LysLeuHisCysAspIleThrProLeuAlaHisAspSerGluAspTyrIly 781
|||||
5 AAGCTCAAACTGACATTAAGTGGTTCACAGAGATTCTGAAGAAGCCGA 54
|||||
781 sleuileGluIntYrLeuLeuAsnThrHisAlaProThrHisLysAspT 798
|||||
55 GATCATCAGGAGATGATGTTAAGAACACTCATGCAACACACACAAATGCGT 104
|||||
798 rpSerLeuGluLeuGluGluValPheSerLeuAspArgAspGlyGluLeu 814
|||||
105 ATCACTTGAAGTCAATCATCTATTAAAGTAGAGCGGTGAAGGCCAATGC 154
|||||
815 AsnLysTyrSerArgTyrLysAsnAsnLeuHisAsnLysMetLeuLeuTr 831
|||||
155 CAGCGTTACAGCCCTTAAG...CAGCTTCATAACCGAAGATTGCTGTG 201
|||||
831 pHisGlySerArgLeuThrAsnPheValGlyIleLeuSerGlnGlyLeuA 848
|||||
202 GCACGGTCCAGGACCACCACTTTGCTGGGATCCTGTCCAGGGTCTTC 251
|||||
848 rgIleAlaProProGluAlaProValThrGlyTyrMetPheGlyLysGly 864
|||||
252 GGATAGCCCGCCCTGAAGCGCCCTGACAGCTACATGTTTGGTAAAGGG 301
|||||
865 LeuTyrPheAlaAspLeuValSerLysSerAlaGlnTyrCysTyrValAs 881
|||||
302 ATCTATTTCGCTGACATGCTCTCAAGAGAGTCCCACTACTGCCATAGTC 351
|||||
881 pArgAsnAsnProValGlyLeuMetLeuLeuSerGluValAlaLeuGlyA 898
|||||
352 TCAGGAGAGACCAATAGCTTAATCTCTTGGGAGAAAGTTGCCCTTGAA 401
|||||
898 sPMeTyrGluLeuLysLysAlaThrSerMetAspLysProProArgGly 914
|||||
402 ACATGTATGAACCTGAAGCAGCTTCACATATCAGCAAGTTACCCAGGGC 451
|||||
915 LysHisSerThrLysGlyLeuGlyLysThrValProLeuGluSerGluPh 931
|||||
452 AAGCAGAGTGTCAAGGTTTGGCAAACTACCCCTGATCCTTCAGCTAA 501
|||||
931 eVallyLysTyrArgAspValValValProCysGlyLysProValProS 948
|||||
502 CATTAGT...CTGGATGGTGTAGAGTTCCTCTTGGGACCGGGATTTCAT 548
|||||
948 eSerileArgSerSerGluLeuMetTyrAsnGluTyrIleValTyrAsn 964
|||||
549 CTGGTGTGAATGACACCTCTCTACTATATACGAGTACATGCTCTATGAT 598
|||||
965 ThrSerGlnValLysMetGlnPheLeuLeuLysValArgPheHisHisLy 981
|||||
599 ATTGCTCAGTAAATCTGAAGTATCTGCTGAACTGAAATCAATTTTAA 648
|||||
981 s 981
649 G 649
seq_name: gb_est1:AV834168
seq_documentation_block:
LOCUS      AV834168      661 bp      mRNA      22-JUN-2001
DEFINITION AV834168 K. Sato unpublished cDNA library: Hordeum vulgare subsp.
            vulgare shoots germination Hordeum vulgare subsp. vulgare cDNA
            clone rbags5K04, mRNA sequence.
ACCESSION  AV834168
VERSION    AV834168.1 GI:14526257
KEYWORDS   EST
SOURCE     Hordeum vulgare subsp. vulgare.
            Hordeum vulgare subsp. vulgare
            Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
            Poaceae; Poideae
            ORGANISM
                ; Triticeae; Hordeum.
                1 (bases 1 to 661)
REFERENCE  Sato, K.
            Barley EST sequencing project in NIG and Okayama Univ
            Unpublished (2001)
            Contact: Kazuhiro Sato
            Research Institute for Bioresources
            Okayama University, Barley Germplasm Center
            Chuo 2-20-1, Kurashiki, Okayama 710-0046, Japan
            Email: kzsato@rib.okayama-u.ac.jp,
            URL: http://www.rib.okayama-u.ac.jp/barley/
            Sato, K., Saitoh, D., Takeda, K., Shini, T. and Kohara, Y. Direct
            submission;
            database: http://www.shigen.nig.ac.jp/barley/Barley.html.
            Location/Qualifiers
                1..661
                /organism="Hordeum vulgare subsp. vulgare"
                /cultivar="Haruna Nijo"
                /db_xref="taxon:112509"
                /clone="rbags5K04"
```

/clone_lib="K. Sato unpublished cDNA library: Hordeum
vulgare subsp. vulgare shoots germination"
/tissue_type="shoots"
/dev_stage="germination"
BASE COUNT 180 a 181 c 132 g 168 t
ORIGIN

alignment_scores:
Quality: 642.00 Length: 136
Ratio: 4.827 Gaps: 1
Percent Similarity: 97.794 Percent Identity: 90.441

alignment_block:

US-09-236-995D-2 x AV834168/rev ..

Align seg 1/1 to reverse of: AV834168 from: 1 to: 661

```

848 ArgfileAlaProGluAlaProValThr.GlyTyrMetPheGlyLysG 864
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
661 AGGATACACCTCTTGAGGACCCGCGGCTATATGTTGGCAAG 612
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
864 lyLeuTyrPheAlaAspLeuValSerLysSerAlaGlnTyrCysTyrVal 880
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
611 GCCTCTACTTTGCAGATTTAGTAAGCAAGAGTGCACAGTATTGTTATGTG 562
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
881 AspArgAsnAsnProValGlyLeuMetLeuLeuSerGluValAlaLeuG1 897
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
561 GATAGAAAATCCGACTGGCTGTGATGCTCTTCTGAGGTTGCTCTAGG 512
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
897 yAspMetTyrGluLeuLysLysAlaThrSerMetAspLysProProArgG 914
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
511 AGCATTTCAGTGAACCAAGCAAGCAAGTGGACCAACCTCCAAGAG 462
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
914 lyLysHisSerThrLysGlyLeuGlyLysThrValProLeuGluSerGlu 930
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
461 GGAAGCATTCGACCAAGGCTTAGCAAAACTGTGCCACTAGAGTCGGAG 412
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
931 PheValLysTyrArgAspValValProCysGlyLysProValPr 947
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
411 TTGTTTAATGAGGAGTATGTCGTCGCTGTCGCAAGCCAGTGCC 362
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
947 oSerSerLeuArgSerSerGluLeuMetTyrAsnGluTyrileValTyrA 964
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
361 AGCATCTATCAGGCGATCTGACCTCTGTACACGAGTATATATGTACA 312
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
964 snThrSerGlnValLysMetGlnPheLeuLeuLysValArgPheHisHis 980
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
311 ACACAGCTCAGTGAAGTGCAGTCTTGTGAAGGTCAAAATTCGGTCAC 262
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
981 LysArg 982
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
261 AAGCGT 256

```

seq_name: gb_est2:BG031594

seq_documentation_block:

LOCUS BG031594 990 bp mRNA EST 24-JAN-2001
DEFINITION 602299739F1 NIH_MGC-87 Homo sapiens cDNA clone IMAGE:4394288 5',
mRNA sequence.

ACCESSION BG031594

VERSION BG031594.1 GI:12422028

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 990)

AUTHORS NIH-MGC <http://mhc.nci.nih.gov/>

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

CONTACT: Robert Strausberg, Ph.D.

Email: cgaabs-r@mail.nih.gov

Tissue Procurement: DCTD/DPD
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM10089 row: m column: 09
High quality sequence stop: 678.

Location/Qualifiers

1..990

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:4394288"

/clone_lib="NIH_MGC_87"

/tissue_type="mammary adenocarcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/note="Organ: breast; Vector: pCMV-SPORT6; Site: 1. NotI;

Site: 2. SalI; Cloned unidirectionally; oligo-gt primed.

Average insert size 1.383 kb. Library enriched for

full-length clones and constructed by Life Technologies.

Note: this is a NIH_MGC Library."

BASE COUNT 274 a 218 c 251 g 247 t

ORIGIN

alignment_scores:

Quality: 642.00 Length: 232

Ratio: 3.326 Gaps: 2

Percent Similarity: 83.190 Percent Identity: 54.310

alignment_block:

US-09-236-995D-2 x BG031594

Align seg 1/1 to: BG031594 from: 1 to: 990

```

750 GlyPheAspSerAspSerAspGluSerLeuAspLysTyrMetLysLe 766
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
9 GGGTCGTGATGATAGCAGCAAGGATCCCATCGATGTCACACTATGAGAAGCT 58
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
766 uHisCysAspIleThrProLeuAlaHisAspSerGluAspTyrLysLeuI 783
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
59 CAAACTGACATTAAAGTGGTGTGACAGAGATCTCAGAGAGCCGAGATCA 108
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
783 leGluGlnTyrLeuLeuAsnThrHisAlaProThrHisLysAspTyrSer 799
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
109 TCAGGAAGTATGTTAAGAACACTCATGCAACACACACAAATGCGTATGAC 158
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
800 LeuGluLeuGluGluValPheSerLeuAspArgAspGlyGluLeuAsnLy 816
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
159 TTGGAAGTCATCGATATCTTTAAGATAGAGCGTGAAGCGGAATGCCAGCG 208
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
816 sTyrSerArgTyrLysAsnAsnLeuHisAsnLysMetLeuLeuTrpHisG 833
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
209 TTACAGCCCTTTAAG...CAGCTTCATAACCCGAAGATTGCTGTGGCAGC 255
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
833 lySerArgLeuThrAsnPheValGlyIleLeuSerGlnGlyLeuArgIle 849
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
256 GGTCCAGGACCAACCACTTCTGCTGGATCTGTCAGGGTCTTCGGATA 305
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
850 AlaProProGluAlaProValThrGlyTyrMetPheGlyLysGlyLeuTy 866
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
306 GCCCGCGCTGAAGCGCCCGTGACAGGCTACATGTTTGGTAAGGATCTA 355
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
866 rPheAlaAspLeuValSerLysSerAlaGlnTyrCysTyrValAspArgA 883
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
356 TTTTCGTCACATGTCCTCCAGAGAGTCCCAACTACTGCCATACGCTCAGG 405
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
883 snAsnProValGlyLeuMetLeuSerGluValAlaLeuGlyAspMet 899
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
406 GAGACCAATAGGCTTAATCTGTTGGAGAACTTCCCTTGGAAACATG 455
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
900 TyrGluLeuLysLysAlaThrSerMetAspLysProProArgGlyLysH 916
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||

```

```

456 TATGAACGAGCAGCCTTCACATATCAGCAAGTTACCAAGGCGCAAGCA 505
916 sSerThrlyGlyLeuGlyLysThrValProLeuGluSerGluPheValL 933
506 CAGTGTCAAAGCTTGGGCAAACTACCCCTGATCCTTCAGCTAACATTA 555
933 yStrpArgaspValValValProCysGlyLysProValProSerSer 949
556 GT...CTGATGGTGTAGACGTTCTCTTGGGACCGGATTCATCTGGT 602
950 IleArgSerSerGluLeuMetTyrAsnGluTyrIleValTyrAsnThrSe 966
603 GTCAATGACACCTCTCTACTATAACAGGTACATGTCTATGATATTGC 652
966 rGlnValLysMetGlnPheLeuLeuLysValArgPheHisLys 981
653 TCAGTA.CATCTGAAGTATCTCTGAAACTTGAATTCAATTGTAAAG 697

seq_name: gb_est2:BI091452
seq_documentation_block: 834 bp mRNA EST 20-JUN-2001
LOCUS 602858931F1 NIH_MGC_10 Homo sapiens cDNA clone IMAGE:5000263 5',
DEFINITION mRNA sequence.
ACCESSION BI091452
VERSION BI091452.1 GI:14509782
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 834)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: c9apbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Incyte Genomics, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1031 row: n column: 08
High quality sequence stop: 804.
FEATURES
SOURCE
1..834
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5000263"
/cell_line="NIH_MGC_10"
/lab_host="DH10B"
/note="Organ: cervix; Vector: pCMV-SPORT6; Site:1: NotI;
Site:2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.5 kb. Library prepared by Life
Technologies."
BASE COUNT 230 a 206 c 214 g 184 t
ORIGIN

```

alignment_scores: Quality: 618.00 Length: 278
Ratio: 2.971 Gaps: 5
Percent Similarity: 74.820 Percent Identity: 47.482

alignment_block:
US-09-236-995D-2 x BI091452
Align seg 1/1 to: BI091452 from: 1 to: 834

seq_name: gb_est2:BG429402

seq_documentation_block: 632 bp mRNA EST

LOCUS BG429402
DEFINITION 602499622F1 NIH_MGC_75 Homo sapiens cDNA clone IMAGE:4613500 5',

14-MAR-2001

/dev_stage="2 days after treatment with beta glucan
elicitator preparation from Phytophthora sojae"
/lab_host="E. coli strain XL0LR"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA
was directionally ligated into the Unizap XR vector from
Stratagene and packaged using Gigapack III Gold packaging
extracts. Plasmids containing cDNA inserts were excised
from the recombinant lambda-zap phage using Ex-assist
helper phage and propagated in XL0LR cells."

BASE COUNT 131 a 87 c 101 g 100 t
ORIGIN

alignment_scores: Quality: 589.00 Length: 139
Ratio: 4.496 Gaps: 0
Percent Similarity: 94.245 Percent Identity: 76.978

alignment_block:

US-09-236-995D-2 x BE941860 ..

Align seg 1/1 to: BE941860 from: 1 to: 419

790 ThrHisAlaProThrHisLysAspTTPSerLeuGluLeuGluValPh 806
|||||
1 ACTCATGCCCCACACATAGGAGTCTTGAGCTAGAGAGAATG 50
806 eSerLeuAspArgAspGlyLeuLeuAsnLysTyrSerArgTyrLysAsnA 823
|||||
51 AGCACTAGAAGAGAGAGGTGAATAAGATCAATATGCGCTTACAGAGACA 100
823 snLeuHisAsnLysMetLeuLeuTrpHisGlySerArgLeuTrpAsnPhe 839
|||||
101 AAGTAGGCAACAGAAATGCTCTTATGGCATGGTCTTAGGCTGACGAACTTT 150
840 ValGlyLeuLeuSerGlnGlyLeuArgIleAlaProProGluAlaProVa 856
|||||
151 GTGGGCATCTTAAACCAAGGACGAGAAATGACCTCCCGAAGCCCTGTC 200
856 lThrGlyTyrMetPheGlyLysGlyLeuTyrPheAlaAspLeuValSerL 873
|||||
201 AACTGGTTATATGTTGGCCCAAGGGATGTACTTTGCTGACCTGTGCAGCA 250
873 ysSerAlaGlnTyrCysTyrValAspArgAsnAsnProValGlyLeuMet 889
|||||
251 AGAGTGTCTAGTATGTTCTCACTGATAAGAAATCTGCTGCTGCTAATG 300
890 LeuLeuSerGluValAlaLeuGlyAspMetTyrGluLeuLysLysAlaTh 906
|||||
301 CTTTGTAGTGAAGTTGCCCTTGGAAATGCTATGAGCTCAAAAAGCTAA 350
906 rSerMetAspLysProProArgGlyLysHisSerThrLysGlyLeuGlyL 923
|||||
351 GTATATGGATAAACCTCCGCAAGGAAGACATTTCTACTAAAGAGACTGGCA 400
923 ysThrValProLeuGlu 928
|||||
401 AGAAATGCCCCCTGGAA 417

seq_name: gb_est2:BG751755

seq_documentation_block:
LOCUS BG751755 771 bp mRNA EST 15-MAY-2001
DEFINITION 602730481F1 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:4874012 5',
mRNA sequence.
ACCESSION BG751755
VERSION BG751755.1 GI:14062408
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS NIH-MGC http://mcc.ncl.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1753 row: a column: 21
High quality sequence stop: 758.

FEATURES
source

1. .771
Location/Qualifiers

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4874012"
/clone_lib="NIH_MGC_43"
/tissue_type="normal pigmented retinal epithelium"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pGB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library. I"

BASE COUNT 221 a 177 c 177 g 196 t
ORIGIN

alignment_scores: Quality: 586.00 Length: 232
Ratio: 3.256 Gaps: 5
Percent Similarity: 77.586 Percent Identity: 50.862

alignment_block:

US-09-236-995D-2 x BG751755 ..

Align seg 1/1 to: BG751755 from: 1 to: 771

762 LysTyrMetLysLeuHisCysAsp..... 769
|||||
50 AAGGCCAGACACCCATTGTGACCAACACATATAGTAGCTCCTCAGATAGA 99
770lThrProLeuAlaHisAspSerGluAspTyrLysLeuLeuGluG 785
|||||
100 GCACCTGGCCGCCCTTGACCATGAAGTTATGAGTTCAAAGTGAATGCC 149
785 lntTyrLeuLeuAsnThrHisAlaProThrHisLysAspTyrSerLeuGlu 801
|||||
150 AGTACCTACATCTACCCATGCTCCACACACAGCAGCTATACCATGACC 199
802 LeuGluGluValPheSerLeuAspArgAspGlyGluLeuAsnLysTyrSe 818
|||||
200 TTGCTGATTTCTTGAAGTGGAGAGGATGCTGAG.....AAAGA 240
818 rArgTyrLysAsnAsnLeuHisAsnLysMetLeuLeuTrpHisGlySerA 835
|||||
241 AGCCTTCAGAGAGGACCTTCATAACAGGATGCTCTATGGCATGGTTCCA 290
835 rglLeuThrAsnPheValGlyIleLeuSerGlnGlyLeuArgIleAlaPro 851
|||||
291 GGATGAGTAACCTGGGTGGGAATCTTGAGCCATGGGCTTCGAATGCCCA 340
852 ProGluAlaProValThrGlyTyrMetPheGlyLysGlyLeuTyrPheAl 868
|||||
341 CCTGAAGCTCCCATCACAGGTTACATGTTTGGGAAAGGAATCTACTTTC 390

